OM of: US-09-801-115-2 to: EST:* out_format : pfs

Date: Jul 19, 2002 8:24 PM

About: Results were produced by the GenCore software, version 4.5,
      Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/ccp2.1/USPTO_spool/US09801115/runat_19072002_182249_6105/app_query.fasta_1.156
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -HEAPSIZE=500
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -THR_MAX=100 -THR_MIN=0
-MINLEN=0 -MAXLEN=200000000 -USER=US09801115_@CGN1_1.2376
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-801-115-2
Query length: 99
Database: EST:*
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 1581.540000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	!	Documentation
gb_est1:AA911088	-	508.00	1174.94	3.4e-56	386	!	AA911088 ok67e01.s1 NCI_CGAP_GC
gb_est1:AI371387	-	508.00	1174.81	3.4e-56	391	!	AI371387 wu2ie10.x1 Soares_Diec
gb_est1:AI989739	-	508.00	1174.68	3.5e-56	396	!	AI989739 wu2ie10.x1 Soares_Diec
gb_est1:AW183193	-	508.00	1174.61	3.5e-56	399	!	AW183193 xj67t11.x1 Soares_NFL
gb_est1:AA516431	-	508.00	1174.53	3.6e-56	402	!	AA516431 ne58a03.s1 NCI_CGAP_CC
gb_est1:AI989747	-	508.00	1174.36	3.6e-56	409	!	AI989747 wu2if09.x1 Soares_NCI_
gb_est1:AI182604	-	508.00	1174.02	3.8e-56	423	!	AI826023 wk35d10.x1 NCI_CGAP_Pr
gb_est1:AI128804	-	508.00	1173.93	3.8e-56	427	!	AI128804 qa94b08.s1 Soares_feta
gb_est1:AA455042	-	508.00	1173.81	3.9e-56	432	!	AA455042 aa04a07.s1 Soares_NhHm
gb_est1:AA429945	-	508.00	1173.36	3.9e-56	432	!	AA429945 zw67f10.s1 Soares_test
gb_est1:AI078580	-	508.00	1173.36	4.0e-56	452	!	AI078580 oz37h05.x1 Soares_NhHm
gb_est1:AI743235	-	508.00	1173.34	4.1e-56	453	!	AI743235 wg90a02.x1 Soares_NSF
gb_est2:BF109912	+	508.00	1173.01	4.3e-56	468	!	BF109912 UI-R-CA1-bjb-b-12-0-UI
gb_est2:BF399486	+	508.00	1172.15	4.8e-56	510	!	BF399486 UI-R-CA1-bjb-b-12-0-UI
gb_est1:BG705303	+	508.00	1171.88	5.0e-56	524	!	BG705303 602d7808f1 NIH_MGC_95
gb_est1:BG031757	+	508.00	1171.69	5.1e-56	534	!	BG031757 602998819F1 NIH_MGC_87
gb_est1:BM472086	+	505.00	1165.59	1.1e-55	487	!	BM472086 AGENCOURT_6469731 NIH
gb_est1:AA989129	-	504.00	1165.58	1.1e-55	386	!	AA989129 or75h10.s1 NCI_CGAP_lu
gb_est1:BG529072	+	504.00	1158.52	2.8e-55	781	!	BG529072 602591699F1 NIH_MGC_60
gb_est2:BG249994	-	504.00	1156.44	3.6e-55	962	!	BG249994 602362062F1 NIH_MGC_90
gb_est2:BF691818	+	491.00	1132.01	1.8e-54	529	!	BF691818 602247787F1 NIH_MGC_52
gb_est2:BG613984	+	484.00	1117.41	5.4e-53	443	!	BG613984 602641513F1 NIH_MGC_61
gb_est2:AV759888	+	480.00	1117.30	1.7e-52	432	!	AV759888 MDS Homo sapi
gb_est1:AI141284	+	476.50	1101.01	4.4e-52	395	!	AI141284 qa52h10.s1 Soares_NhHm
gb_est1:AI265924	+	475.00	1098.62	6.0e-52	353	!	AI265924 qx68d08.s1 NCI_CGAP_Ov
gb_est1:BF203359	+	475.00	1098.92	2.1e-51	930	!	BF203359 601865968F1 NIH_MGC_17
gb_est1:W52820	+	472.00	1088.78	2.1e-51	468	!	W52820 zc55c06.r1 Soares_senec
gb_est2:AA876539	-	471.50	1086.42	2.9e-51	527	!	AA876539 oe48h08.s1 NCI_CGAP_Pr
gb_est2:AA702011	-	471.50	1085.98	3.0e-51	551	!	AA702011 zi70d02.s1 Soares_feta
gb_htc:AF151058	-	471.50	1083.69	4.1e-51	692	!	AF151058 Homo sapiens HSPC224 m
gb_est2:AV683453	-	471.50	1082.55	4.7e-51	776	!	AV683453 Homo sapi
gb_est2:N89912	-	471.00	1087.79	2.4e-51	409	!	N89912 zb22g09.s1 Soares_fetal
gb_est2:BE737159	-	467.50	1073.02	1.6e-50	789	!	BE737159 601305091F1 NIH_MGC_39
gb_est2:BG532587	+	463.00	1066.23	3.8e-50	543	!	BG532587 602542152F1 NIH_MGC_61
gb_est1:BI561562	+	460.50	1058.16	1.1e-49	678	!	BI561562 603256264F1 NIH_MGC_97
gb_est1:AA987264	-	452.00	1045.39	5.5e-49	333	!	AA987264 os11a06.s1 NCI_CGAP_lu
gb_est1:AI636679	-	445.50	1026.11	6.6e-48	500	!	AI636679 th71c04.x1 Soares_NhHm
gb_est1:BF791773	-	445.50	1023.42	9.3e-48	654	!	BF791773 602251973F1 NIH_MGC_84
gb_est2:AI263261	-	444.00	1026.89	6.0e-48	326	!	AI263261 qp62b12.x1 Soares_feta
gb_est2:W93646	-	442.50	1017.17	2.1e-47	606	!	W93646 zq97d00.s1 Soares_fetal
gb_est2:BE875161	+	436.00	1004.81	1.0e-46	456	!	BE875161 601488641F1 NIH_MGC_69

seq_name: gb_est1:AA911088

seq_documentation_block:
LOCUS	AA911088		386 bp	mRNA	linear

DEFINITION ok67e01.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1519032 3',
        mRNA sequence.
ACCESSION   AA911088
VERSION   AA911088.1 GI:3050378
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 386)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
          Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
          Emmert-Buck, M.D., Ph.D.
          cDNA Library Preparation: M. Bento Soares, Ph.D.
          cDNA Library Arrayed by: Greg Lennon, Ph.D.
          DNA Sequencing by: Washington University Genome Sequencing Center
          Clone distribution: NCI-CGAP clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          www-bio.llnl.gov/bbrp/image/image.html
          Insert Length: 381 Std Error: 0.00
          Seq primer: -40m13 fwd. ET from Amersham
          High quality sequence stop: 79.
FEATURES   Location/Qualifiers
     source   1..386
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:1519032"
                 /clone_lib="NCI_CGAP_GC4"
                 /tissue_type="pooled germ cell tumors"
                 /lab_host="DH10B"
                 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                 polylinker; 1st strand cDNA was prepared from 3 pooled
                 germ cell tumors, and was then primed with a Not I -
                 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                 RI adaptors (Pharmacia), digested with Not I and cloned
                 into the Not I and Eco RI sites of the modified pT7T3
                 vector. Library is normalized. Library was constructed by
                 Bento Soares and M. Fatima Bonaldo."

BASE COUNT	113 a	87 c	81 g	105 t

ORIGIN

alignment_scores:
    Quality: 508.00   Length: 99
    Ratio: 5.131   Gaps: 0
    Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AA911088/rev

Align seg 1/1 to reverse of: AA911088 from: 1 to: 386

    1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVal 17
      |||||||||||||||||||||||||||||||||||||||||||||||||||
  357 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGT 308

   17 LysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
      |||||||||||||||||||||||||||||||||||||||||||||||||

Left column

```
307 GAAAGGCCACGTGAAGAATGCTGCGCGCTGGATATTATCAACTCACTGGTAA 258
 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGlyThr 50
257 CAACAGTATTCATCGTCATCGTATCTGTGTTGGCACTGATACCAGAAACC 208
207 ACAACATTGACACAGTTGGTGGAGGGGTGTTTGCACTTGTTGGACAGCAGTATG 158
 51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy 67
207 ... 
 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnPros 84
157 CTGTCTTGCCGACGGCGGGGCCCTTATTTACCGGAAGCTTCGTTCAATCCCA 108
 84 erGlyProTyrGlnLysLysProValHisGlyLysLysGluLysLysGluValLeu 99
107 GCGGTCCTTACCAGAAAAAAGCCTGTGCATGAAAAAAAGAGAAGTTTTG 61

seq_name: gb_est1:AI989739

seq_documentation_block:
LOCUS       AI989739                391 bp    mRNA    linear    EST 27-OCT-1999
DEFINITION  wu2ie10.x1 Soares_Dieckgraefe_colon_NHCD Homo sapiens cDNA clone
            IMAGE:2520714 3', mRNA sequence.
ACCESSION   AI989739
VERSION     AI989739.1  GI:5836620
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 391)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1..391
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Soares_Dieckgraefe_colon_NHCD"
                     /clone="IMAGE:2520714"
                     /tissue_type="colonic mucosa from 3 patients with Crohn's
                     disease"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     This clone is available royalty-free through LLNL; contact the
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Tissue samples
                     provided by Dr. Brian Dieckgraefe (Washington University,
                     dieck@im.wustl.edu); colonic mucosa represents a range of
                     disease involvement from moderate to severe Crohn's
                     disease; samples include both perforating (fistulas) and
                     non-perforating samples. Library constructed by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT      113 a     90 c     81 g    107 t
ORIGIN

alignment_scores:
  Quality:        508.00      Length:        99
  Ratio:            5.131      Gaps:           0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

Right column

```
alignment_block:
  US-09-801-115-2 x AI989739/rev .

Align seg 1/1 to reverse of: AI989739  from: 1 to 391
```

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
358 ATGGATAACGTGCAGCCGAAATAAAAACATCGGCCCTTCTCGTTCAGTGT 309
 17 LLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
308 GAAAGGCCACGTGAAGAATGCTGCGCGCTGGATATTATCAACTCACTGGTAA 259
 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGlyThr 50
258 CAACAGTATTCATCGTCATCGTATCTGTGTTGGCACTGATACCAGAAACC 209
 51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy 67
208 ACAACATTGACACAGTTGGTGGAGGGGTGTTTGCACTTGTTGACAGCAGTATG 159
 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnPros 84
158 CTGTCTTGCCGACGGCGGGGCCCTTATTTACCGGAAGCTTCGTTCAATCCCA 109
 84 erGlyProTyrGlnLysLysProValHisGlyLysLysGluLysLysGluValLeu 99
108 GCGGTCCTTACCAGAAAAAAGCCTGTGCATGAAAAAAAGAGAAGTTTTG 62

seq_name: gb_est1:AW183193

seq_documentation_block:
LOCUS       AW183193                396 bp    mRNA    linear    EST 18-NOV-1999
DEFINITION  xj67f11.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2662317 3', mRNA sequence.
ACCESSION   AW183193
VERSION     AW183193.1  GI:6451679
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 396)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1..396
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2662317"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729006-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT      115 a     92 c     80 g    109 t
ORIGIN
```

Left column

```
alignment_scores:
    Quality:     508.00       Length:       99
    Ratio:       5.131        Gaps:          0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AW183193/rev

Align seg 1/1  to reverse of: AW183193  from: 1 to: 396

1   MetAspAsnValGlnProLysIleLeuLysHisArgProPheCysPheSerVa  17
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
359 ATGGATAACGTGCAGCCCAAAATAAAACATCGCCCCTTCTGTTCAGTGT  310

17  lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT  34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
309 GAAAGGCCACGTGAAGATGCTCGGGCTGGATAATTATCAACTCACTGGTAA  260

34  hrThrValPheMetLeuIleValSerValAlaAlaLeuIleProGluThr  50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
259 CAACAGTATTCATGCTCATCGTATCGTGTTGGCACTGGTCACCAGAAACC  210

51  ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy  67
    |||||||||||||||||||||||||||||||||||||||||||||||
209 ACAACATTGACACAGTTGGGTGGAGGGGTGTTTGCACTTGTCACCAGTATG  160

67  sCysLeuAlaAspGlyAlaLeuIleLeuTyrArgLysLeuLeuPheAsnPros  84
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
159 CTGTGCTTGCCGACGGGGCCCTTAATTACCGGAAGCTTCTTGTTCAATCCA  110

84  erGlyProGlnTyrGlnLeuLysProValHisGlyLysLysGluValLeu  99
    ||||||||||||||||||||||||||||||||||||||||||||||||||
109 GCGGTCTTACCAGAAAAAAGCCTGTGCATGAAAAAAAAGAAGTTTTG  63
```

```
seq_documentation_block:
seq_name: gb_est1.AA516431

LOCUS       AA516431    399 bp    mRNA    linear    EST 13-AUG-1997
DEFINITION  ne58a03.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:901516 3',
            mRNA sequence.
ACCESSION   AA516431
VERSION     AA516431.1  GI:2255955
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 399)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.
            , Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arraying: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 527   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1..399
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:901516"
                     /clone_lib="NCI_CGAP_Co3"
                     /sex="pooled"
```

Right column

```
BASE COUNT    115 a   94 c   81 g   109 t
ORIGIN
```

```
                     /tissue_type="colon"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker: Site_1; Not I; Site_2; Eco RI; 1st strand cDNA
                     was prepared from 12 pooled bulk tumor samples and primed
                     with a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector. Library went through one round of
                     normalization."

alignment_block:
US-09-801-115-2 x AA516431/rev

Align seg 1/1  to reverse of: AA516431  ..

alignment_scores:
    Quality:     508.00       Length:       99
    Ratio:       5.131        Gaps:          0
Percent Similarity: 100.000    Percent Identity: 100.000

seq_name: gb_est1.AI989747

Align seg 1/1  to reverse of: AA516431  from: 1 to: 399

1   MetAspAsnValGlnProLysIleLeuLysHisArgProPheCysPheSerVa  17
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
357 ATGGATAACGTGCAGCCCAAAATAAAACATCGCCCCTTCTGTTCAGTGT  308

17  lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT  34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
307 GAAAGGCCACGTGAAGATGCTCGGGCTGGATAATTATCAACTCACTGGTAA  258

34  hrThrValPheMetLeuIleValSerValAlaAlaLeuIleProGluThr  50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
257 CAACAGTATTCATGCTCATCGTATCGTGTTGGCACTGGTCACCAGAAACC  208

51  ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy  67
    |||||||||||||||||||||||||||||||||||||||||||||||
207 ACAACATTGACACAGTTGGGTGGAGGGGTGTTTGCACTTGTCACCAGTATG  158

67  sCysLeuAlaAspGlyAlaLeuIleLeuTyrArgLysLeuLeuPheAsnPros  84
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
157 CTGTGCTTGCCGACGGGGCCCTTAATTACCGGAAGCTTCTTGTTCAATCCA  108

84  erGlyProGlnTyrGlnLeuLysProValHisGlyLysLysGluValLeu  99
    ||||||||||||||||||||||||||||||||||||||||||||||||||
107 GCGGTCTTACCAGAAAAAAGCCTGTGCATGAAAAAAAAGAAGTTTTG  61
```

```
seq_documentation_block:
seq_name: gb_est1.AI989747

LOCUS       AI989747    402 bp    mRNA    linear    EST 27-OCT-1999
DEFINITION  wu21f09.x1 Soares_Dieckgraefe_colon_NHCD Homo sapiens cDNA clone
            IMAGE:2520713 3', mRNA sequence.
ACCESSION   AI989747
VERSION     AI989747.1  GI:5836628
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 402)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -400P from Gibco.
FEATURES             Location/Qualifiers
     source          1..402
```

```
REFERENCE   1  (bases 1 to 409)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  SOURCE    human.
  KEYWORDS  EST.
  VERSION   AI826623.1  GI:5447294
  ACCESSION AI826623
  DEFINITION wk35010.x1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:2417395 3',
            mRNA sequence.
  LOCUS     AI826623              409 bp    mRNA    linear   EST 21-DEC-1999

seq_documentation_block:

seq_name: gb_est1:AI826623
```

```
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:2520713"
  /clone_lib="Soares_Dieckgraefe_colon_NHCD"
  /tissue_type="colonic mucosa from 3 patients with Crohn's
  disease"
  /lab_host="DH10B (phage-resistant)"
  /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
  modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
  strand cDNA was primed with a Not I - oligo(dT) primer [5'
  TGTTACCAATCTGAAGTGGGAGGACGACGCCCGGGCCGCNNNNTTTTTTTTTTTTTTTT 3'],
  double-stranded cDNA was ligated to Eco RI adaptors
  (Pharmacia), digested with Not I and cloned into the Not I
  and Eco RI sites of the modified pT7T3 vector. Library
  went through one round of normalization. Tissue samples
  provided by Dr. Brian Dieckgraefe (Washington University,
  dieck@im.wustl.edu); colonic mucosa represents a range of
  disease involvement from moderate to severe Crohn's
  disease; samples include both perforating (fistulas) and
  non-perforating samples. Library constructed by Bento
  Soares and M. Fatima Bonaldo.
  "
BASE COUNT      113 a     90 c     82 g    117 t
ORIGIN

alignment_scores:
        Quality:   508.00            Length:    99
        Ratio:       5.131           Gaps:       0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AI989747/rev

Align seq 1/1  to reverse of: AI989747   from: 1  to 402

  1 MetAspAsnValGlnProLysIleLeuLysHisArgProPheCysPheSerVa 17
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
368 ATGGATAACGTGCAGCCGAAAATAAAACACTCGCCCTTCTGCTTCAGTGT 319

 17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
318 GAAAGGCCACGTGAAGATGCTGCGGCTGGATAATTATCAACTCACTGGTAA 269

 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGlyThr 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
268 CAACAGTATTCAATGCTCATCGTATCTGTGGTGCACTGATACCAGAAACC 219

 51 ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy 67
    |||||||||||||||||||||||||||||||||||||||||||||||
218 ACAACAATTGACACGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATG 169

 67 sCysLeuLeuAspGlyAlaLeuIleTyrArgLysLeuPheAsnPros 84
    |||||||||||||||||||||||||||||||||||||||||||||||
168 CTGTCTTGCCGACGGGCGCCCTTATTTACCGGAAGCTTCGTTCAATCCA 119

 84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
    |||||||||||||||||||||||||||||||||||||||||||||||
118 GCGGTCCTTACCAGAAAAAAGCCTGTCGCATGAAAAAAAAGAAGTTTG 72

seq_name: gb_est1:AI826623
```

```
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE     Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Michael J. Brownstein, M.D., Ph.D.
          Tissue Preparation: M. Bento Soares, Ph.D.
                              Emmett-Buck, M.D., Ph.D.
          cDNA Library Arrayed by: Greg Lennon, Ph.D.
          cDNA Library Preparation: Michael J. Brownstein, M.D., Ph.D., Michael R.
          DNA Sequencing by: Washington University Genome Sequencing Center
          Clone distribution: NCI-CGAP clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          www-bio.llnl.gov/bbrp/image/image.html
          Insert Length: 462   Std Error: 0.00
          Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
          source          1..409
                          /organism="Homo sapiens"
                          /db_xref="taxon:9606"
                          /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                          with a modified polylinker; 1st strand cDNA was prepared
                          from normal prostate bulk tissue, and was then primed with
                          a Not I - oligo(dT) primer. Double-stranded cDNA was
                          ligated to Eco RI adaptors (Pharmacia), digested with Not
                          I and cloned into the Not I and Eco RI sites of the
                          modified pT7T3 vector. Library is normalized, and was
                          constructed by Bento Soares and M. Fatima Bonaldo.
                          "
                          /lab_host="DH10B"
                          /clone_lib="NCI_CGAP_Pr22"
                          /clone="IMAGE:2417395"
                          /tissue_type="normal prostate"
                          /sex="male"
BASE COUNT      112 a     94 c     80 g    123 t
ORIGIN

alignment_block:
US-09-801-115-2 x AI826623/rev

alignment_scores:
        Quality:   508.00            Length:    99
        Ratio:       5.131           Gaps:       0
Percent Similarity: 100.000   Percent Identity: 100.000

Align seq 1/1  to reverse of: AI826623   from: 1  to 409

  1 MetAspAsnValGlnProLysIleLeuLysHisArgProPheCysPheSerVa 17
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
367 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGT 318

 17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
317 GAAAGGCCACGTGAAGATGCTGCGGCTGGATAATTATCAACTCACTGGTAA 268

 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGlyThr 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
267 CAACAGTATTCAATGCTCATCGTATCTGTGGTGCACTGATACCAGAAACC 218

 51 ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy 67
    |||||||||||||||||||||||||||||||||||||||||||||||
217 ACAACAATTGACACGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATG 168

 67 sCysLeuLeuAspGlyAlaLeuIleTyrArgLysLeuPheAsnPros 84
    |||||||||||||||||||||||||||||||||||||||||||||||
167 CTGTCTTGCCGACGGGCGCCCTTATTTACCGGAAGCTTCGTTCAATCCA 118

 84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
    |||||||||||||||||||||||||||||||||||||||||||||||
117 GCGGTCCTTACCAGAAAAAAGCCTGTCGCATGAAAAAAAAGAAGTTTG 71

seq_name: gb_est1:AI128804
```

```
seq_documentation_block:
LOCUS       AI128804              423 bp    mRNA    linear   EST 05-OCT-1998
DEFINITION  ga94b08.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:1694391 3', mRNA sequence.
ACCESSION   AI128804
VERSION     AI128804.1  GI:3597318
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 382   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 413.
FEATURES             Location/Qualifiers
     source          1..423
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1694391"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGGCGCGCCANTCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
BASE COUNT      116 a      101 c      89 g     117 t
ORIGIN

alignment_scores:
        Quality:        508.00          Length:      99
        Ratio:          5.131           Gaps:         0
Percent Similarity: 100.000       Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AI128804/rev

Align seg 1/1  to reverse of: AI128804  from: 1  to: 423

1   MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa  17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
353 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTTCTGCTTCAGTGT  304

17  lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT  34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
303 GAAAGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA  254

34  hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr  50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
253 CAACAGTATTCATCCTCAACGTAACTGTGTTGGCACTGATACCAGAAACC  204

51  ThrThrLeuThrValGlyGlyValGlyValAlaPheAlaLeuValThrAlaValGly  67
    ||||||||||||||||||||||||||||||||||||||||||||||||
203 ACAACAGTTGACAGTTGGTGTGGAGGGGTTGTTGGTGTTTGCACTGGTGAACAGCAGTATG 154
```

```
seq_name: gb_est1:AA455042

seq_documentation_block:
LOCUS       AA455042              427 bp    mRNA    linear   EST 06-JUN-1997
DEFINITION  aa04a07.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:812244
            3', mRNA sequence.
ACCESSION   AA455042
VERSION     AA455042.1  GI:2177818
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 427)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            , T., Waterston,R. and Wilson,R.
  TITLE     Washu-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 395.
FEATURES             Location/Qualifiers
     source          1..427
                     /organism="Homo sapiens"
                     /db_xref="GDB:6043155"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:812244"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      119 a      101 c      91 g     116 t
ORIGIN

alignment_scores:
        Quality:        508.00          Length:      99
        Ratio:          5.131           Gaps:         0
Percent Similarity: 100.000       Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AA455042/rev

Align seg 1/1  to reverse of: AA455042  from: 1  to: 427

67  sCysLeuAlaAspGlyAlaLeuLeuIleTyrArgLysLeuLeuPheAsnProS  84
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
153 CTGTCTTGCCGACGGGGCGCCCTTTATTTACCGGAAGCTTCTGTTCAATCCA  104

84  erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu  99
    |||||||||||||||||||||||||||||||||||||||||||||||
103 GCCGTCGTCCTTACCAGAAAAAAGCCTGTGCATGAAAAAAAGAGAGTTTTG  57
```

seq_name: gb_est1:AA429945

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
356 ATGGATAACGGTGCAACGCCGAAAATAAAACATCGGCCCTTCTGCTTCAGTGT 307

 17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
306 GAAAGGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 257

 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuLeuIleProGluThr 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
256 CAACAGTATTCATGCTCATCGTATCGTGTGTTGGCACTGATACCAGAAACC 207

 51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy 67
    ||||||||||||||||||||||||||||||||||||||||||||||||||
206 ACAACAGTTGACAGTTGGTGTGGAGGGGTGTTTGCACTTGTGACACAGTATG 157

 67 sCysLeuAlaAspGlyAlaLeuLeuIleTyrArgLysLeuLeuPheAsnProS 84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
156 CTGTCTTGCCGACGGGCGCCCTTATTTACCGGAAGCTTCGTTCAATCCA 107

 84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
    ||||||||||||||||||||||||||||||||||||||||||||||||||
106 GCGGTCCTTACCAGACAAAAAGCCTGTGCATGAAAAAAGAAGAGTTTTG 60
```

seq_documentation_block:
```
LOCUS       AA429945               432 bp    mRNA    linear   EST 16-OCT-1997
DEFINITION  zw67f10.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:781291
            3', mRNA sequence.
ACCESSION   AA429945
VERSION     AA429945.1  GI:2113244
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 432)
AUTHORS     Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
TITLE       Washu-Merck EST Project 1997
JOURNAL     Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 421.
FEATURES             Location/Qualifiers
     source          1..432
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:781291"
                     /sex="male"
                     /clone_lib="Soares_testis_NHT"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc., and primed with a Not I / oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCCAATTTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
```

seq_name: gb_est1:AI078580

```
BASE COUNT     119 a    101 c    92 g    120 t
ORIGIN

alignment_scores:
        Quality:         508.00          Length:          99
        Ratio:             5.131         Gaps:             0
        Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AA429945/rev

Align seg 1/1 to reverse of: AA429945    from: 1 to: 432

  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
361 ATGGATAACGGTGCAACGCCGAAAATAAAACATCGGCCCTTCTGCTTCAGTGT 312

 17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
311 GAAAGGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 262

 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuLeuIleProGluThr 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
261 CAACAGTATTCATGCTCATCGTATCGTGTGTTGGCACTGATACCAGAAACC 212

 51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy 67
    ||||||||||||||||||||||||||||||||||||||||||||||||||
211 ACAACAGTTGACAGTTGGTGTGGAGGGGTGTTTGCACTTGTGACACAGTATG 162

 67 sCysLeuAlaAspGlyAlaLeuLeuIleTyrArgLysLeuLeuPheAsnProS 84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
161 CTGTCTTGCCGACGGGCGCCCTTATTTACCGGAAGCTTCGTTCAATCCA 112

 84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
    ||||||||||||||||||||||||||||||||||||||||||||||||||
111 GCGGTCCTTACCAGACAAAAAGCCTGTGCATGAAAAAAGAAGAGTTTTG 65
```

seq_documentation_block:
```
LOCUS       AI078580               452 bp    mRNA    linear   EST 10-AUG-1998
DEFINITION  oz37h05.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1677561
            3', mRNA sequence.
ACCESSION   AI078580
VERSION     AI078580.1  GI:3412988
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 452)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 410.
FEATURES             Location/Qualifiers
     source          1..452
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1677561"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
```

Site: 2; Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NbHPU, and fetal heart NbHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT   125 a   110 c   97 g   120 t
ORIGIN

alignment_scores:
  Quality: 508.00   Length: 99
  Ratio: 5.131   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AI078580/rev

Align seq 1/1 to reverse of: AI078580   from: 1 to: 452

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
359 ATGGATAACGTGCCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGT 310
 17 LeuGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
309 GAAAGGCCACGTGAAGAAGTGCCGGCTGGATATTATCAACTCACTGGTAA 260
 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
259 CAACAGTATTCATGCTCATCGTATCCTGTGTTGGCACTGATACCAGAAACC 210
 51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy 67
209 ACAACATTGACACGTGGTGGAGGGGTGTTGCACTTGTGACAGCAGTATG 160
 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnPros 84
159 CTGTCTTGCCGACGGGCGCCCTTATTTACCGGAAGCTTCTGTTCAATCCA 110
 84 erGlyProTyrGlnLysLysProValHisGlyLysLysGluValLeu 99
109 GCGGTCCTTACCAGAAAAAAGCCTGTGCATGAAAAAAAAGAAGTTTTG 63
```

seq_name: gb_est1:AI743235

seq_documentation_block:
LOCUS   AI743235   453 bp   mRNA   linear   EST 20-DEC-1999
DEFINITION wg90a02.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:2372330 3', mRNA sequence.
ACCESSION   AI743235
VERSION   AI743235.1   GI:5111523
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 453)
  AUTHORS
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco.
            Insert Length: 378   Std Error: 0.00
FEATURES   Location/Qualifiers
     source   1..453

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2372330"
/clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NbHSF pool 1: 309384-310919, 323208-325895 Soares Nb2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares Nb2HR8-9W pool 1: 758080-760583, 772104-774407 Soares NbHPA pool 1: 304776-306311, 320136-332823, 326280-326663 Soares NbHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT   123 a   109 c   93 g   128 t
ORIGIN

alignment_scores:
  Quality: 508.00   Length: 99
  Ratio: 5.131   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AI743235/rev

Align seq 1/1 to reverse of: AI743235   from: 1 to: 453

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
362 ATGGATAACGTGCCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGT 313
 17 LeuGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
312 GAAAGGCCACGTGAAGAAGTGCCGGCTGGATATTATCAACTCACTGGTAA 263
 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
262 CAACAGTATTCATGCTCATCGTATCCTGTGTTGGCACTGATACCAGAAACC 213
 51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy 67
212 ACAACATTGACACGTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATG 163
 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnPros 84
162 CTGTCTTGCCGACGGGCGCCCTTATTTACCGGAAGCTTCTGTTCAATCCA 113
 84 erGlyProTyrGlnLysLysProValHisGlyLysLysGluValLeu 99
112 GCGGTCCTTACCAGAAAAAAGCCTGTGCATGAAAAAAAAGAAGTTTTG 66
```

seq_name: gb_est2:BF109912

seq_documentation_block:
LOCUS   BF109912   468 bp   mRNA   linear   EST 20-OCT-2000
DEFINITION 7i71f03.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:3526805 3' similar to TR:Q9UI41 Q9UI41 CHEMOKINE-LIKE FACTOR;
            , mRNA sequence.
ACCESSION   BF109912
VERSION   BF109912.1   GI:10939602
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 468)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Seq primer: -40UP from Gibco.
FEATURES
     source
             1..468
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="IMAGE:3526805"
             /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
             /lab_host="DH10B"
             /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
             a modified polylinker; Site_1: Not I; Site_2: Eco RI;
             Equal amounts of plasmid DNA from five normalized
             libraries were mixed, and ss circles were made in vitro.
             Following HAP purification, this DNA was used as tracer in
             a subtractive hybridization reaction. The driver was
             PCR-amplified cDNAs from pools of 5,000 clones made from
             the same 5 libraries. The pools consisted of the following
             libraries and cloneIds: Soares NbHSF pool 1:
             309384-310919, 323208-325895 Soares Nb2HP pool 1:
             145033-147335, 147720-148103, 148872-149255, 15002 -
             150407, 151176-152327 Soares Nb2HF8 9W pool 1:
             758280-760583, 772104-774407 Soares NbHPA pool 1:
             304776-306311, 320136-322823, 326280-326663 Soares NbHQT
             pool 1: 723720-726407, 739080-740999 Subtraction by Bento
             Soares and M. Fatima Bonaldo."
BASE COUNT   127 a   116 c   102 g   123 t
ORIGIN

alignment_scores:
   Quality: 508.00   Length: 99
   Ratio: 5.131   Gaps: 0
   Percent Similarity: 100.000   Percent Identity: 100.000

Align seq 1/1 to reverse of: BF109912 from: 1 to: 468
US-09-801-115-2 x BF109912/rev   ..

alignment_block:

   1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa   17
 362 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCGTTCAGTGT   313
  17 LLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT   34
 312 GAAAGGCCACGTGAAGATGCTCGGGCTGGATATTATCAACTCACTGGTAA   263
  34 hrThrValPheMetLeuIleLeuValSerValLeuAlaLeuLeuProGluThr 50
 262 CAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATTACCAGAAACC 213
  51 ThrThrLeuThrValGlyGlyValPheAlaLeuValAlaValThrAlaValCy 67
 212 ACAACAGTTGACAGTTGGTTGGAGGGGTGTTTGCACTTGTTGCACCAGTATG 163
  67 sCysLeuAlaAspGlyAlaLeuLeuTyrArgLysLeuLeuProHisAsnProS 84
 162 CTGTCTTGCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCA   113
  84 erGlyProTyrGlnLysTyrGlnLysLysProValHisGluLysLysGluValLeu 99
 112 GCGGTCCTTACCAGAAAGAAAAAGCCTGTGCATGAAAAAAAGAAGTTTTG   66

seq_name: gb_est2:BF399486

---

seq_documentation_block:
LOCUS   BF399486   510 bp   mRNA   linear   EST 27-NOV-2000
DEFINITION UI-R-CA1-bjb-b-12-0-UI.s1 UI-R-CA1 Rattus norvegicus cDNA clone
            UI-R-CA1-bjb-b-12-0-UI 3', mRNA sequence.
ACCESSION   BF399486
VERSION   BF399486.1 GI:11384494
KEYWORDS   EST.
SOURCE   Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 510)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE   Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
COMMENT   Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            clones will be available through Research Genetics (www.resgen.com)
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES
     source
            1..510
            /organism="Rattus norvegicus"
            /strain="Sprague-Dawley"
            /db_xref="taxon:10116"
            /clone="UI-R-CA1-bjb-b-12-0-UI"
            /clone_lib="UI-R-CA1"
            /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
            polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CA1
            library is a subtracted library derived from the following
            tissues: thalamus, cerebellum, hypothalamus, medulla, pons
            , midbrain, cerebral cortex, corpus striatum, testis, and
            hippocampus. For a detailed description of the library
            from which this clone was derived, please visit our web
            site at ratest.eng.uiowa.edu. The subtraction has been
            previously described in (Bonaldo, Lennon and Soares,
            Genome Research 6:791-806, 1996)
            TAG_SEQ="None found"
BASE COUNT   130 a   131 c   104 g   145 t
ORIGIN

alignment_scores:
   Quality: 508.00   Length: 99
   Ratio: 5.131   Gaps: 0
   Percent Similarity: 100.000   Percent Identity: 100.000

Align seq 1/1 to reverse of: BF399486 from: 1 to: 510
US-09-801-115-2 x BF399486/rev   ..

alignment_block:

   1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa   17
 371 ATGGATAACGTGCAGCCCAAAATAAAACATCGCCCCTTCTGCTTCAGTGT   322
  17 LLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT   34
 321 GAAAGGCCACGTGAAGATGCTCCGGCTGGATATTATCAACTCACTGGTAA   272

Left column

```
  34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr  50
     |||||||||||||||||||||||||||||||||||||||||||||||||
 271 CAACAGTATTCATCTCATCGTATCTGTGTGTGGCCACTGATACCAGAAACC 222

  51 ThrThrLeuThrValGlyGlyValGlyValPheAlaLeuValAlaValCy  67
     ||||||||||||||||||||||||||||||||||||||||||||||||
 221 ACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTGGTGGCACAGTATG  172

  67 sCysLeuLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnPros  84
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 171 CTGTCTTGCCGACGGGCGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCA  122

  84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu  99
     |||||||||||||||||||||||||||||||||||||||||||||||
 121 GCGGTCCTTACCAGAAAAAAGCCTGTGCATGAAAAAAAGAAGTTTTG   75
```

seq_name: gb_est2:BG705303

seq_documentation_block:
```
LOCUS       BG705303               524 bp    mRNA    linear   EST 07-MAY-2001
DEFINITION  602687808F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4820568 5',
            mRNA sequence.
ACCESSION   BG705303
VERSION     BG705303.1  GI:13979504
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 524)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10725 row: o column: 01
            High quality sequence stop: 521.
FEATURES             Location/Qualifiers
     source          1..524
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4820568"
                     /clone_lib="NIH_MGC_95"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescript (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTTTTVN-3',
                     size-selected for average insert size 2.5 kb and
                     normalized to ROt 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation). Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      142 a    113 c    138 g    131 t
ORIGIN
```

alignment_block:
US-09-801-115-2 x BG705303

alignment_scores:
```
        Quality:           508.00    Length:                 99
        Ratio:               5.131   Gaps:                    0
        Percent Similarity: 100.000  Percent Identity:  100.000
```

Right column

Align seq 1/1 to: BG705303 from: 1 to: 524

```
   1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa  17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 141 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTTCTGCTTCAGTGT 190

  17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT  34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 191 GAAAGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 240

  34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr  50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 241 CAACAGTATTCATCTCATCGTATCTGTGTGTGGCCACTGATACCAGAAACC 290

  51 ThrThrLeuThrValGlyGlyValGlyValPheAlaLeuValAlaValCy  67
     ||||||||||||||||||||||||||||||||||||||||||||||||
 291 ACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTGGTGGCACAGTATG  340

  67 sCysLeuLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnPros  84
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 341 CTGTCTTGCCGACGGGCGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCA  390

  84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu  99
     |||||||||||||||||||||||||||||||||||||||||||||||
 391 GCGGTCCTTACCAGAAAAAAGCCTGTGCATGAAAAAAAGAAGTTTTG 437
```

seq_name: gb_est2:BG031757

seq_documentation_block:
```
LOCUS       BG031757               534 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  602998189F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4394093 5',
            mRNA sequence.
ACCESSION   BG031757
VERSION     BG031757.1  GI:12422364
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 534)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10089 row: e column: 06
            High quality sequence stop: 446.
FEATURES             Location/Qualifiers
     source          1..534
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4394093"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      163 a    108 c    132 g    131 t
ORIGIN
```

alignment_scores:

```
                        Quality:   508.00       Length:       99
                          Ratio:   5.131          Gaps:        0
Percent Similarity: 100.000          Percent Identity: 100.000

alignment_block:
us-09-801-115-2 x BG031757

Align seg 1/1  to: BG031757    from: 1  to: 534

         us-09-801-115-2  .  .

   1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 123 ATGGATAACGTGCAGCCCAAAAATAAAACATCGCCCTTCTGCTTCAGTGT 172

  17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 173 GAAAGGCCACGTGAAAGAATGCTCGGCTGGATATTATCAAACTCACTGGTAA 222

  34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 223 CAACAGTATTCAATGCTCATCGTATCCGTGTTGGCACTCGATACCAGAAACC 272

  51 ThrThrLeuThrValGlyLyGlyValPheAlaLeuValThrAlaValCy 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 273 ACAACAATTGACACAGTTGGTGGGAGGGGTGTTTGCACTTGACAGCAGTATG 322

  67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnProS 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 323 CTGTCTTGCCGACGGGGCGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCA 372

  84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
     |||||||||||||||||||||||||||||||||||||||||||||||
 373 GCCGGTCCTTACCAGAAAAAGCCTGTGCATGAAAAAAGAAGTTTTG 419
```

OM of: US-09-801-115-2 to: N_Geneseq_032802:*   out_format : pfs

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Date: Jul 19, 2002 8:28 PM

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09801115/runat_19072002_182249_6141/app_query.fasta_1.156
 -DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
 -QGAPOP=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000000
 -USER=US09801115_@CGN1_1_170 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-801-115-2
Query length: 99
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 176.440000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT/NA1999.DAT:AAX97873	+	508.00	1190.77	5.6e-58	41	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT/NA1998.DAT:AAV59598	+	508.00	1188.67	7.3e-58	50	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT/NA1998.DAT:AAV59746	+	508.00	1188.63	7.4e-58	50	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAA87730	+	508.00	1188.44	7.5e-58	51	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAF64012	+	508.00	1188.44	7.5e-58	51	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAA38006	-	508.00	1188.06	7.9e-58	53	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAH98515	+	508.00	1187.98	8.0e-58	50	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAH98548	-	508.00	1187.98	8.0e-58	50	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAH34435	+	508.00	1187.59	8.4e-58	50	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAS44932	+	508.00	1187.50	8.5e-58	50	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAA15919	+	504.00	1184.76	1.2e-57	29	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAA15929	+	504.00	1119.98	2.2e-57	46	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAX97826	+	489.00	1143.66	2.4e-55	50	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAA38007	+	471.50	1103.03	4.3e-53	45	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAS56747	+	471.50	1099.57	6.7e-53	63	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT/NA1999.DAT:AAZ34051	+	471.50	1099.14	7.1e-53	66	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAZ34052	+	471.50	1099.14	7.1e-53	66	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAC58238	+	471.50	1099.14	7.1e-53	66	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAA87771	+	471.50	1099.05	7.2e-53	66	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAF64053	+	471.50	1099.05	7.2e-53	66	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAC00147	+	461.00	1079.17	9.2e-52	41	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAA16090	+	461.00	1075.82	1.1e-52	39	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAS44120	+	357.00	832.94	4.8e-38	39	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAS45121	+	339.50	792.22	8.6e-36	36	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAS45120	+	339.50	787.53	1.6e-35	20	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAS45121	+	339.50	787.53	1.6e-35	20	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAA38008	+	331.00	778.14	5.4e-35	32	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT/NA1999.DAT:AAX97651	+	323.00	754.36	1.1e-33	32	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAZ34052	+	318.50	739.10	8.1e-33	49	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAF78499	+	318.50	739.10	8.1e-33	49	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAF93358	+	318.50	739.10	8.1e-33	49	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT/NA1999.DAT:AAX41509	+	317.00	739.09	8.1e-33	4	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:ABA09082	-	294.00	683.78	9.7e-30	35	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAK09082	-	257.50	594.99	8.6e-25	37	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAK68359	+	215.00	474.58	4.4e-18	2	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAK68360	+	215.00	474.58	4.4e-18	2	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAK68361	+	215.00	474.58	4.4e-18	1	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAK82406	+	211.00	448.84	1.2e-16	1	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAK82406	-	211.00	448.84	1.2e-16	1	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAC90012	-	170.00	379.79	8.3e-13	9	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT/NA2001A.DAT:AAK82407	+	142.50	293.80	5.1e-08	6	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT/NA2000.DAT:AAC06850	+	141.50	325.46	8.8e-10	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT/NA1999.DAT:AAA16224	+	116.50	254.62	7.8e-06	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:AAS61351	-	108.00	212.65	0.0017	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:AAS61352	+	102.00	198.41	0.0105	
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAH98802	+	88.00	186.40	0.0492	

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:AAX97873

seq_documentation_block:
ID   AAX97873 standard; cDNA; 413 BP.
XX
AC   AAX97873;
XX
DT   23-SEP-1999 (first entry)
XX
DE   Human secreted protein encoding cDNA #61.
XX
KW   Secreted protein; human; cytostatic; thrombotic; osteopathic; forensic;
KW   diagnostic; gene therapy; chromosome mapping; secretion vector; ss.
XX
OS   Homo sapiens.
XX
PN   WO9258825-A2.
XX
PD   27-MAY-1999.
XX
PF   13-NOV-1998;   98WO-IB01862.
XX
PR   13-NOV-1998;   98US-0099273;
PR   04-SEP-1998;   98US-0099273;
PR   13-NOV-1997;   97US-0066677;
PR   17-DEC-1997;   97US-0069957;
PR   09-FEB-1998;   98US-0074121;
PR   13-APR-1998;   98US-0081563;
PR   10-AUG-1998;   98US-0096116.
XX
PA   (GEST) GENSET.
XX
PI   Bouguelert L, Duclert A, Dumas Milne Edwards J;
XX
PT   Extended cDNAs encoding secreted proteins
PS   Claim 1; Page 261; 307pp; English.
XX
CC   AAX97813-X97906 represent extended cDNA's which encode novel human
CC   secreted proteins (see Aar36129-Y362222) and which have cytostatic,
CC   thrombotic and osteopathic activity. The extended cDNAs can be used to
CC   express secreted proteins or parts of them or to obtain antibodies
CC   capable of binding to the secreted proteins. They may also be used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   Uses also include design of expression vectors and secretion vectors.
XX
SQ   Sequence 413 BP; 121 A; 82 C; 95 G; 115 T; 0 other;

DR   WPI; 1999-347472/29.
DR   P-PSDB; AAY36189.

alignment_scores:
         Quality: 508.00   Length: 99
         Ratio: 5.131   Gaps: 0
         Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AAX97873

Align seq 1/1 to: AAX97873   from: 1 to: 413

     1 MetAspAsnValGlnProLysHisArgProPheCysPheSerVa 17
        |||||||||||||||||||||||||||||||||||||||||||||||||||
     1 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGT 95

    17 LeuLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
        |||||||||||||||||||||||||||||||||||||||||||||||||||
    46 ... 95

```
seq_name: /SIDS1/gcgdata/geneseqn-embl/NA1998.DAT:AAV59598

seq_documentation_block:
ID   AAV59598 standard; DNA; 504 BP.
XX
AC   AAV59598;
XX
DT   06-JAN-1999  (first entry)
XX
DE   Human secreted protein gene 88 clone HAUAV32.
XX
OS   Homo sapiens.
XX
PN   WO983944B-A2.
XX
PD   11-SEP-1998.
XX
PF   06-MAR-1998;  98WO-US04493.
XX
PR   02-OCT-1997;  97US-0061060.
PR   07-MAR-1997;  97US-0038621.
PR   07-MAR-1997;  97US-0040161.
PR   07-MAR-1997;  97US-0040162.
PR   07-MAR-1997;  97US-0040163.
PR   07-MAR-1997;  97US-0040333.
PR   07-MAR-1997;  97US-0040334.
PR   07-MAR-1997;  97US-0040336.
PR   07-MAR-1997;  97US-0040626.
PR   11-APR-1997;  97US-0043311.
PR   11-APR-1997;  97US-0043312.
PR   11-APR-1997;  97US-0043313.
PR   11-APR-1997;  97US-0043314.
PR   11-APR-1997;  97US-0043568.
PR   11-APR-1997;  97US-0043569.
PR   11-APR-1997;  97US-0043576.
PR   11-APR-1997;  97US-0043578.
PR   11-APR-1997;  97US-0043580.
PR   11-APR-1997;  97US-0043669.
PR   11-APR-1997;  97US-0043670.
PR   11-APR-1997;  97US-0043671.
PR   11-APR-1997;  97US-0043672.
PR   11-APR-1997;  97US-0043674.
PR   23-MAY-1997;  97US-0047492.
PR   23-MAY-1997;  97US-0047500.
PR   23-MAY-1997;  97US-0047501.
PR   23-MAY-1997;  97US-0047502.

PR   23-MAY-1997;  97US-0047503.
PR   23-MAY-1997;  97US-0047581.
PR   23-MAY-1997;  97US-0047582.
PR   23-MAY-1997;  97US-0047583.
PR   23-MAY-1997;  97US-0047584.
PR   23-MAY-1997;  97US-0047585.
PR   23-MAY-1997;  97US-0047586.
PR   23-MAY-1997;  97US-0047587.
PR   23-MAY-1997;  97US-0047588.
PR   23-MAY-1997;  97US-0047589.
PR   23-MAY-1997;  97US-0047590.
PR   23-MAY-1997;  97US-0047592.
PR   23-MAY-1997;  97US-0047593.
PR   23-MAY-1997;  97US-0047594.
PR   23-MAY-1997;  97US-0047595.
PR   23-MAY-1997;  97US-0047596.
PR   23-MAY-1997;  97US-0047597.
PR   23-MAY-1997;  97US-0047598.
PR   23-MAY-1997;  97US-0047599.
PR   23-MAY-1997;  97US-0047600.
PR   23-MAY-1997;  97US-0047601.
PR   23-MAY-1997;  97US-0047612.
PR   23-MAY-1997;  97US-0047613.
PR   23-MAY-1997;  97US-0047614.
PR   23-MAY-1997;  97US-0047615.
PR   23-MAY-1997;  97US-0047617.
PR   23-MAY-1997;  97US-0047618.
PR   23-MAY-1997;  97US-0047633.
PR   06-JUN-1997;  97US-0048964.
PR   06-JUN-1997;  97US-0048974.
PR   05-JUN-1997;  97US-0049610.
PR   13-JUN-1997;  97US-0049926.
PR   08-JUL-1997;  97US-0051926.
PR   16-JUL-1997;  97US-0055724.
PR   18-AUG-1997;  97US-0055630.
PR   22-AUG-1997;  97US-0056631.
PR   22-AUG-1997;  97US-0056632.
PR   22-AUG-1997;  97US-0056636.
PR   22-AUG-1997;  97US-0056637.
PR   22-AUG-1997;  97US-0056662.
PR   22-AUG-1997;  97US-0056664.
PR   22-AUG-1997;  97US-0056845.
PR   22-AUG-1997;  97US-0056862.
PR   22-AUG-1997;  97US-0056864.
PR   22-AUG-1997;  97US-0056872.
PR   22-AUG-1997;  97US-0056874.
PR   22-AUG-1997;  97US-0056875.
PR   22-AUG-1997;  97US-0056876.
PR   22-AUG-1997;  97US-0056877.
PR   22-AUG-1997;  97US-0056878.
PR   22-AUG-1997;  97US-0056879.
PR   22-AUG-1997;  97US-0056880.
PR   22-AUG-1997;  97US-0056881.
PR   22-AUG-1997;  97US-0056882.
PR   22-AUG-1997;  97US-0056884.
PR   22-AUG-1997;  97US-0056886.
PR   22-AUG-1997;  97US-0056887.
PR   22-AUG-1997;  97US-0056888.
PR   22-AUG-1997;  97US-0056889.
PR   22-AUG-1997;  97US-0056892.
PR   22-AUG-1997;  97US-0056893.
PR   22-AUG-1997;  97US-0056894.
PR   22-AUG-1997;  97US-0056903.
PR   22-AUG-1997;  97US-0056908.
PR   22-AUG-1997;  97US-0056909.
PR   22-AUG-1997;  97US-0056910.
PR   22-AUG-1997;  97US-0056911.
PR   05-SEP-1997;  97US-0057650.
PR   05-SEP-1997;  97US-0057669.
PR   05-SEP-1997;  97US-0057761.
PR   12-SEP-1997;  97US-0058785.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
```

```
 96  GAAAGGCCACGTGAAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA  145
 34  hrThrValPheMetLeuIleValLeuSerValLeuAlaLeuIleProGluThr  50
146  CAACAGTATTCAATCCTCATCGTATCTGTTGTTGGCACTGATACCAGAAAAC  195
 51  ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy  67
     |||||||||||||||||||||||||||||||||||||||||
196  ACAACATTGACAGTTGGTGGAGGGGGTGTTTTGCACTTGTGACAGCAGTATG  245
 67  sCysLeuAlaAspGlyAlaLeuIleThrTyrArgLysLeuLeuPheAsnProS  84
     |||||||||||||||||||||||||||||||||||
246  CTGCCTTGCCGACGGGGCCCTTATTACCTACAGAAAAAGCTTCTGTCAATCCA  295
 84  erGlyProTyrGlnLeuLysProValHisGlyLysLysGluValLeu  99
     |||||||||||||||||||||||||||||||||||
296  GCGGTCCTTACCAGAAAAAGCCTGTCATGAAAAAAAAGAAGTTTG  342
```

```
ID   AAV59746 standard; DNA; 506 BP.
seq_documentation_block;
seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV59746
```

CC   This sequence represents a nucleic acid molecule designated Gene 88 from
CC   the human cDNA clone HAUAV32 (deposited as clone ATCC 97901 and ATCC
CC   209047) which encodes a secreted human protein. The gene can be used to
CC   generate fusion proteins by linking to the gene to a human
CC   immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of
CC   the fused protein as compared to the human protein only.
CC   The invention relates to 186 novel genes and their fragments (nucleic
CC   acid sequences AAV59511-V59812; amino acid sequences AAW74731-W75026)
CC   which are useful for preventing, treating or ameliorating medical
CC   conditions e.g. by protein or gene therapy. Also, pathological conditions
CC   can be diagnosed by determining the amount of the new polypeptides in a
CC   sample or by determining the presence of mutations in the new
CC   polynucleotides. Specific uses are described for each of the 186 new
CC   polynucleotides, based on which tissues they are most highly expressed in
CC   (see AAV59511 for described uses).
XX
SQ   Sequence 504 BP; 143 A; 106 C; 128 G; 127 T; 0 other;

```
alignment_scores:
      Quality:       508.00              Length:          99
      Ratio:           5.131             Gaps:             0
Percent Similarity: 100.000      Percent Identity: 100.000 .

alignment_block;
US-09-801-115-2 x AAV59598

Align seq 1/1   to: AAV59598      ..

            1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
              ||||||||||||||||||||||||||||||||||||||||||||||||||
          131 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGT 180

           17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
              ||||||||||||||||||||||||||||||||||||||||||||||||||
          181 GAAAGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 230

           34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
              ||||||||||||||||||||||||||||||||||||||||||||||||||
          231 CAACAGTATTCATGCTAATCGTATCGTGTTGGCACTGATACCAGAAACC 280

           51 ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy 67
              ||||||||||||||||||||||||||||||||||||||||||||||||
          281 ACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACACAGTATG 330

           67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnPros 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||
          331 CTGGTCTTGCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCA 380

           84 erGlyProTyrGlnLysLysProValHisGlyLeuLysGluValLeu 99
              |||||||||||||||||||||||||||||||||||||||||||||||
          381 GCGGTCCTTACCAGAAAAAGCCTGTCCATGAAAAAAGAAGTTTTG 427
```

CC   The invention relates to 186 novel genes and their fragments (nucleic
CC   acid sequences: AAV59511-V59812; amino acid sequences AAW74731-W75026)
CC   which are useful for preventing, treating or ameliorating medical
CC   conditions e.g. by protein or gene therapy. Also, pathological
CC   conditions can be diagnosed by determining the amount of the new
CC   polypeptides in a sample or by determining the presence of mutations in
CC   the new polynucleotides. Specific uses are described for each of the 186
CC   polynucleotides, based on which tissues they are most highly expressed in
CC   (see AAV59511 for described uses).
XX
SQ   Sequence 506 BP; 148 A; 103 C; 121 G; 132 T; 2 other;

alignment_scores:
            Quality: 508.00   Length: 99
            Ratio: 5.131   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AAV59746

Align seq 1/1 to: AAV59746   from: 1 to: 506

   1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 117 ATGGATAACGTGCAGCCCAAAATAAAACATCGCCCTTTCTGCTTCAGTGT 166

  17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 167 GAAAGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 216

  34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 217 CAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACC 266

  51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuAlaLeuValThrAlaValCy 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 267 ACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGCAGCAGTATG 316

  67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnProS 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 317 CTGTCTTGCCGACGGGCGCCCTTATTTACCGGAAGCTTCGTTCAATCCA 366

  84 erGlyProTyrGlnLysGlySProValHisSGlnLysLysGluValLeu 99
     ||||||||||||||||||||||||||||||||||||||||||||
 367 GCGGTCCTTACCAGAAAAAGCCTGTGCATGAAAAAAAGAAGTTTG 413

seq_name: /SIDS1/gcgdata/geneseqn-embl/NA2000.DAT:AAA87730
ID   AAA87730 standard; cDNA; 515 BP.

seq_documentation_block:
XX
AC   AAA87730;
XX
DT   28-NOV-2000   (first entry)
XX
DE   Human secreted protein encoding cDNA SEQ ID #29.
XX
KW   Human; secreted protein; forensic procedure; gene therapy;
KW   chromosome mapping; cancer; autoimmune disease; cardiovascular disorder;
KW   cystic fibrosis; hypothyroidism; immunological disorder; amyloidosis;
KW   brain disorder; skeletal muscle disorder; eye disorder; obesity;
KW   mitochondriocytopathy; diabetes; atherosclerosis; Alzheimer's disease;
KW   neurodegenerative disorder; graft rejection; dementia; hyperlipidaemia;
KW   septic shock; impotence; ss.
XX
OS   Homo sapiens.
XX
PN   WO20003749l-A2.
XX
PD   29-JUN-2000.
XX
PF   20-DEC-1999;   99WO-IB02058.

PR 23-MAY-1997;   97US-0047613.
PR 23-MAY-1997;   97US-0047614.
PR 23-MAY-1997;   97US-0047615.
PR 23-MAY-1997;   97US-0047617.
PR 23-MAY-1997;   97US-0047618.
PR 23-MAY-1997;   97US-0047632.
PR 23-MAY-1997;   97US-0047633.
PR 06-JUN-1997;   97US-0048964.
PR 06-JUN-1997;   97US-0048974.
PR 13-JUN-1997;   97US-0049610.
PR 08-JUL-1997;   97US-0051926.
PR 16-JUL-1997;   97US-0052874.
PR 18-AUG-1997;   97US-0055724.
PR 22-AUG-1997;   97US-0056630.
PR 22-AUG-1997;   97US-0056631.
PR 22-AUG-1997;   97US-0056632.
PR 22-AUG-1997;   97US-0056636.
PR 22-AUG-1997;   97US-0056637.
PR 22-AUG-1997;   97US-0056662.
PR 22-AUG-1997;   97US-0056664.
PR 22-AUG-1997;   97US-0056845.
PR 22-AUG-1997;   97US-0056862.
PR 22-AUG-1997;   97US-0056864.
PR 22-AUG-1997;   97US-0056872.
PR 22-AUG-1997;   97US-0056874.
PR 22-AUG-1997;   97US-0056875.
PR 22-AUG-1997;   97US-0056876.
PR 22-AUG-1997;   97US-0056877.
PR 22-AUG-1997;   97US-0056878.
PR 22-AUG-1997;   97US-0056879.
PR 22-AUG-1997;   97US-0056880.
PR 22-AUG-1997;   97US-0056881.
PR 22-AUG-1997;   97US-0056882.
PR 22-AUG-1997;   97US-0056884.
PR 22-AUG-1997;   97US-0056886.
PR 22-AUG-1997;   97US-0056887.
PR 22-AUG-1997;   97US-0056888.
PR 22-AUG-1997;   97US-0056889.
PR 22-AUG-1997;   97US-0056892.
PR 22-AUG-1997;   97US-0056893.
PR 22-AUG-1997;   97US-0056894.
PR 22-AUG-1997;   97US-0056903.
PR 22-AUG-1997;   97US-0056908.
PR 22-AUG-1997;   97US-0056909.
PR 22-AUG-1997;   97US-0056910.
PR 22-AUG-1997;   97US-0056911.
PR 05-SEP-1997;   97US-0057650.
PR 05-SEP-1997;   97US-0057669.
PR 22-AUG-1997;   97US-0057761.
PR 12-SEP-1997;   97US-0058785.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
DR WPI: 1998-506364/43.
DR P-PSDB; AAW74961.
XX
Claim 1; Page 472; 721pp; English.

PS   This sequence represents a nucleic acid molecule designated Gene 88 from
CC   the human cDNA clone HAUAV32 (deposited as clone ATCC 97897 and ATCC
CC   209043) which encodes a secreted human protein. The gene can be used to
CC   generate fusion proteins by linking to the gene to a human
CC   immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of
CC   the fused protein as compared to the human protein only.

New isolated human genes and the secreted polypeptide(s) they encode
- useful for diagnosis and treatment of e.g. cancers, neurological
disorders, immune diseases, inflammation or blood disorders

Left column

```
XX
PR   22-DEC-1998;   98US-0113686;
PR   25-JUN-1999;   99US-0141032.
XX
PA   (GEST ) GENSET.
XX
PI   Bougueleret L,  Dumas J,  Duclert A;
XX
DR   WPI; 2000-442637/38.
DR   P-PSDB; AAB25768.
XX
DE   Polynucleotides and polypeptides encoding proteins with signal
DE   peptides, useful in diagnostic, forensic, gene therapy and chromosome
DE   mapping procedures -
XX
PT   Claim 1; Page 169-170; 306pp; English.
XX
```

CC   This sequence represents human cDNA encoding a secreted protein. The
invention relates to sequences AAB7725-A87774 which encode human
secreted proteins AAB25763-B25812. The proteins include signal peptides.
Included in the invention are a host cell containing one of the cDNA
sequences, and a purified antibody capable of binding to one of the
secreted proteins. Also contained in the invention are methods for
storing the sequence data on a computer system, and a method for
identifying features of the cDNA sequences using a computer programme.
The cDNAs are useful for expressing secreted proteins or fragments to
obtain antibodies capable of specifically binding to the secreted
proteins. The cDNAs may also be useful in diagnostic, forensic, gene
therapy and chromosome mapping procedures and may be used to design
expression vectors and secretion vectors. The proteins of the invention
may be used to treat diseases including cancer, autoimmune diseases,
cardiovascular disorders, cystic fibrosis, hypothyroidism, immunological
disorders, amyloidosis, brain disorders, skeletal muscle disorders, eye
disorders, obesity, mitochondriocytopathies, diabetes, atherosclerosis,
neurodegenerative disorders, graft rejection, Alzheimer's disease,
dementia, hyperlipidaemia, septic shock and impotence.

```
SQ   Sequence 515 BP; 143 A; 106 C; 135 G; 130 T; 1 other;
XX
```

alignment_scores:
```
                  Quality:   508.00          Length:        99
                  Ratio:       5.131         Gaps:           0
Percent Similarity: 100.000       Percent Identity: 100.000
```

alignment_block:
US-09-801-115-2 x AAA87730

Align seg 1/1 to: AAA87730   from: 1 to: 515   ..

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa  17
144 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCGTTCAGTGT    193
    |||||||||||||||||||||||||||||||||||||||||||||||
144 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCGTTCAGTGT    193

 17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT  34
194 GAAAGGCCACGTGAAGARGCTGCGGCTGGATATTATCAACTCACTGGTAA  243
    ||||||||||||||||||||||||||||||||||||||||||||||||||
194 GAAAGGCCACGTGAAGARGCTGCGGCTGGATATTATCAACTCACTGGTAA  243

 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr  50
244 CAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGACTACCAGAAACC 293
    ||||||||||||||||||||||||||||||||||||||||||||||||||
244 CAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGACTACCAGAAACC 293

 51 ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy     67
294 ACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATG  343
    ||||||||||||||||||||||||||||||||||||||||||||||||||
294 ACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATG  343

 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnProS  84
344 CTGTCTTGCCGACGGGGCGCCCTTATTTACCGGAAGCTTCGTTCAATCCA  393
    ||||||||||||||||||||||||||||||||||||||||||||||||||
344 CTGTCTTGCCGACGGGGCGCCCTTATTTACCGGAAGCTTCGTTCAATCCA  393

 84 erGlyProTyrGlnLysLysLysProValHisGluLysLysGluValLeu  99
394 GCGGTCCTTACCAGAAAAAAAAGCCTGTGCATGAAAAAAAAGAAGAGTTTTG 440
    |||||||||||||||||||||||||||||||||||||||||||||||||||
394 GCGGTCCTTACCAGAAAAAAAAGCCTGTGCATGAAAAAAAAGAAGAGTTTTG 440
```

Right column

```
seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAF64012
XX
seq_documentation_block:
ID   AAF64012 standard; cDNA; 515 BP.
XX
AC   AAF64012;
XX
DT   05-APR-2001 (first entry)
XX
DE   cDNA encoding human secreted protein #13.
XX
KW   Secreted protein; prevention; treatment; diagnosis; disease;
KW   infection; ds.
XX
OS   Homo sapiens.
XX
PN   WO200100806-A2.
XX
PD   04-JAN-2001.
XX
PF   21-JUN-2000; 2000WO-IB00951.
XX
PR   25-JUN-1999;   99US-0141032.
PR   21-DEC-1999;   99US-0469099.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Bougueleret L,  Jobert S;
XX
DR   WPI; 2001-071487/08.
XX
PT   Claim 1; Page 225; 307pp; English.
XX
```

CC   The present invention relates to 49 secreted proteins and the cDNAs
encoding them. The protein and nucleic acids may be used in the
prevention, treatment and diagnosis of diseases associated with
inappropriate protein expression.

```
SQ   Sequence 515 BP; 143 A; 106 C; 135 G; 130 T; 1 other;
XX
```

alignment_scores:
```
                  Quality:   508.00          Length:        99
                  Ratio:       5.131         Gaps:           0
Percent Similarity: 100.000       Percent Identity: 100.000
```

alignment_block:
US-09-801-115-2 x AAF64012

Align seg 1/1 to: AAF64012   from: 1 to: 515   ..

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa  17
144 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCGTTCAGTGT    193
    |||||||||||||||||||||||||||||||||||||||||||||||
144 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCGTTCAGTGT    193

 17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT  34
194 GAAAGGCCACGTGAAGARGCTGCGGCTGGATATTATCAACTCACTGGTAA  243
    ||||||||||||||||||||||||||||||||||||||||||||||||||
194 GAAAGGCCACGTGAAGARGCTGCGGCTGGATATTATCAACTCACTGGTAA  243

 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr  50
244 CAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGACTACCAGAAACC 293
    ||||||||||||||||||||||||||||||||||||||||||||||||||
244 CAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGACTACCAGAAACC 293

 51 ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy     67
294 ACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATG  343
    ||||||||||||||||||||||||||||||||||||||||||||||||||
294 ACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATG  343

 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnProS  84
344 ...
```

344 CTGTCTTGCCGACGGGGCCCCTTATTTACCGGAAGCTTCTGTTCAATCCCA 393
    |||||||||||||||||||||||||||||||||||||||||||||||||||
394 GCCGGTCCTTACCAGAAAAAGCCTGTGCATGAAAAAAAAGAAGTTTTG 440

84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
   |||||||||||||||||||||||||||||||||||||||||||||||

seq_name: /SIDS1/gcgdata/geneseqn-embl/NA2000.DAT:AAA38006 .

seq_documentation_block:
ID AAA38006 standard; cDNA; 534 BP.
XX
AC AAA38006;
XX
DT 22-AUG-2000 (first entry)
XX
DE UCK-1 nucleotide sequence.
XX
KW UCK-1; chemotaxis factor; immunocyte; haemopoiesis stimulant; tumour;
KW radiotherapy; chemotherapy; human; ss.
XX
OS Homo sapiens.
XX
PN CN1244584-A.
XX
PD 16-FEB-2000.
XX
PF 14-MAY-1999;   99CN-0107284.
PR 14-MAY-1999;   99CN-0107284.
PA (UYBE-) UNIV BEIJING MEDICAL.
XX
PI Ma D, Han W, Zhang Y;
XX
DR WPI; 2000-388170/34.
DR P-PSDB; AAY98142.
XX
PT Chemotaxic factor useful for treatment and diagnosis of immunocyte
PT disorders - has immunocyte chemotaxic stimulating factor
XX
PS Example 4; Fig 1; 31pp; Chinese.
XX
CC This sequence represents an UCK-1 cDNA sequence encoding a chemotaxis
CC factor polypeptide. The UCK-1 protein exhibits immunocyte chemotaxis
CC activity and a haemopoiesis stimulating effect. The invention relates to
CC UCK proteins, their encoding nucleotide sequences and antibodies and
CC antagonists against the proteins. The nucleotide and protein sequences
CC are useful for the preparation of a composition for the diagnosis and
CC treatment of diseases associated with abnormal immunocyte function and
CC low haemopoiesis function caused by radiotherapy and chemotherapy used to
CC treat tumours and other diseases.
XX
SQ Sequence 534 BP; 156 A; 110 C; 134 G; 134 T; 0 other;

alignment_scores:
          Quality: 508.00   Length: 99
          Ratio: 5.131   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AAA38006

Align seq 1/1 to: AAA38006   from: 1 to: 534

  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17

 17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
152 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGT 201

 17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
202 GAAAGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 251

---

34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
   ||||||||||||||||||||||||||||||||||||||||||||||||||
252 CAACAGTATTCATGCTCATCGTATCCGTGTTGGCACTGATACCAGAAACC 301

51 ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy 67
   |||||||||||||||||||||||||||||||||||||||||||||||
302 ACAACATTGACACAGTTGGTGGGAGGGGTGTTTGCACTGTGTGACAGCATG 351

84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
   |||||||||||||||||||||||||||||||||||||||||||||||
402 GCCGGTCCTTACCAGAAAAAGCCTGTGCATGAAAAAAAAGAAGTTTTG 448

67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnPros 84
   ||||||||||||||||||||||||||||||||||||||||||||||||||
352 CTGTCTTGCCGACGGGGCCCCTTATTTACCGGAAGCTTCTGTTCAATCCCA 401

seq_name: /SIDS1/gcgdata/geneseqn-embl/NA2001A.DAT:AAH98515

seq_documentation_block:
ID AAH98515 standard; cDNA; 538 BP.
XX
AC AAH98515;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST-derived coding sequence SEQ ID NO: 372.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition; ss.
XX
OS Homo sapiens.
XX
PN WO20015477-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02687.
XX
PR 25-JAN-2000; 2000US-0491404.
PR 17-JUL-2000; 2000US-0617746.
PR 03-AUG-2000; 2000US-0631451.
PR 15-SEP-2000; 2000US-0663870.
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX
DR WPI; 2001-476164/51.
DR P-PSDB; AAM23856.
XX
PT Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
XX
PS Claim 1; Page 451; 1275pp; English.
XX
CC The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a cDNA
CC of the invention.
XX
SQ Sequence 538 BP; 141 A; 142 C; 115 G; 140 T; 0 other;

alignment_scores:
          Quality: 508.00   Length: 99
          Ratio: 5.131   Gaps: 0

Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AAH98515/rev   ..

Align seg 1/1 to reverse of: AAH98515 from: 1 to: 538

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
366 ATGGATAACGTCCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGT 317

 17 LLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
316 GAAAGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 267

 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
266 CAACAGTATTCAATGCTCATCGTATTCGTGTTTGGCACTGATACCAGAAACC 217

 51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy 67
    ||||||||||||||||||||||||||||||||||||||||||||||||||
216 ACAACAATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATG 167

 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLlysLeuLeuPheAsnProS 84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
166 CTGTCTTGCCGACGGGGCGCCCTTATTTACCGGAAGCTTCTGTTCAATCCA 117

 84 erGlyProTyrGlnLysLysProValHisGluLysLysGlyValLeu 99
    ||||||||||||||||||||||||||||||||||||||||||||||
116 GCGGTCCTTACCAGAAAAAAAGCCTGTGCATGAAAAAAAGAAGAGTTTTG 70
```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAH98548

seq_documentation_block:
ID AAH98548 standard; cDNA; 538 BP.
XX
AC AAH98548;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST-derived coding sequence SEQ ID NO: 405.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition; ss.
XX
OS Homo sapiens.
XX
PN WO20015477-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02687.
XX
PR 25-JAN-2000; 2000US-0491404.
PR 17-JUL-2000; 2000US-0617746.
PR 03-AUG-2000; 2000US-0631451.
PR 15-SEP-2000; 2000US-0663870.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX
DR WPI; 2001-476164/51.
DR P-PSD; AAM23889.
XX
PT Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
XX
PS Claim 1; Page 467; 1275pp; English.
XX

CC The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a cDNA
CC of the invention.
XX
SQ Sequence 538 BP; 141 A; 142 C; 115 G; 140 T; 0 other;

alignment_scores:
Quality: 508.00   Length: 99
Ratio: 5.131   Gaps: 0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AAH98548/rev   ..

Align seg 1/1 to reverse of: AAH98548 from: 1 to: 538

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
366 ATGGATAACGTCCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGT 317

 17 LLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
316 GAAAGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 267

 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
266 CAACAGTATTCAATGCTCATCGTATTCGTGTTTGGCACTGATACCAGAAACC 217

 51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy 67
    ||||||||||||||||||||||||||||||||||||||||||||||||||
216 ACAACAATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATG 167

 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnProS 84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
166 CTGTCTTGCCGACGGGGCGCCCTTATTTACCGGAAGCTTCTGTTCAATCCA 117

 84 erGlyProTyrGlnLysLysProValHisGluLysLysGlyValLeu 99
    ||||||||||||||||||||||||||||||||||||||||||||||
116 GCGGTCCTTACCAGAAAAAAAGCCTGTGCATGAAAAAAAGAAGAGTTTTG 70
```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAH34835

seq_documentation_block:
ID AAH34835 standard; cDNA; 558 BP.
XX
AC AAH34835;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:1917.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX

PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
XX
DR P-PSDB; AAG75430.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 1; Page 3428; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome associated with the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 558 BP; 150 A; 119 C; 145 G; 141 T; 3 other;

alignment_scores:
        Quality: 508.00
        Ratio: 5.131
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-801-115-2 x AAH34835

Align seq 1/1   to: AAH34835 . from: 1   to: 558

```
    1 MetAspAsnValGlnProLysIleLeuHisHisArgProPheCysPheSerVa 17
  147 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGT 196

   17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
  197 GAAAGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 246

   34 hrThrValPheMetLeuLeuIleValSerValLeuAlaLeuIleProGluThr 50
  247 CAACAGTATTCATGCTCATCCATCGTTGTTGGCCACTGATACCAGAAACC 296

   51 ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCys 67
  297 ACAACATTCACAGTTGGTGGAGGGGTGTTTGCACTTGTACAGCAGTATG 346

   67 sCysLeuAlaAspGlyAlaLeuIleThrArgLysLeuuPheAsnProS 84
  347 CTGTCTTGCCGACGGGGCCCTTATTACCGGAAGCTTCTGTTCAATCCA 396

   84 erGlyProTyrGlnLysProValHisGlyLysLysGluValLeu 99
  397 GCGGTCCTTACCAGAAAAAGCCTGTGCATGAAAAAAAGAAGTTTG 443
```

seq_name: /SIDS1/gcgdata/geneseq/geneseqp-embl/NA2001A.DAT:AAS44932

seq_documentation_block:
ID AAS44932 standard; cDNA; 655 BP.
XX
AC AAS44932;

XX
DT 18-DEC-2001 (first entry)
XX
DE cDNA encoding novel human secretory protein, Seq ID No 13.
XX
KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen; ss.
XX
OS Homo sapiens.
XX
PN WO200166689-A2.
XX
PD 13-SEP-2001.
XX
PF 05-MAR-2001; 2001WO-US04942.
XX
PR 07-MAR-2000; 2000US-0519705.
PR 19-MAY-2000; 2000US-0574454.
PR 17-JUN-2000; 2000US-0596193.
PR 14-JUL-2000; 2000US-0616847.
PR 19-SEP-2000; 2000US-0665363.
PR 20-OCT-2000; 2000US-0693267.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
DR WPI; 2001-589934/66.
DR P-PSDB; AAU28032.
XX
PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS Claim 1; SEQ ID No 13; 107pp; English.
XX
CC The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides

CC   analgesic effects or other pain reducing effects, immunoglobulin like
CC   activity and can act as an antigen in a vaccine composition to raise an
CC   immune response. AAS44920-AAS45295 represent novel human secreted protein
CC   coding sequences of the invention.
XX
SQ   Sequence 655 BP; 178 A; 129 C; 181 G; 167 T; 0 other;

alignment_scores:
           Quality: 508.00
           Ratio: 5.131
Percent Similarity: 100.000   Percent Identity: 100.000.

alignment_block:
US-09-801-115-2 x AAS44932

Align seg 1/1 to: AAS44932 from: 1 to: 655

       1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     134 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGT 183

      17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     184 GAAAGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 233

      34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     234 CAACAGTATTCAtGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACC 283

      51 ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy 67
         |||||||||||||||||||||||||||||||||||||||||||||||
     284 ACAACATTGACACAGTTGGTGGAGGGGTGTTTGCACTGTGACAGCAGTATG 333

      67 sCysLeuAlaAspGlyAlaLeuIleTyrArgArgLysLeuPheAsnPros 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     334 CTGTCTTGCCGACGGGCGCCTTATTACCGGAAGCTTCTGTTCAATCCA 383

      84 erGlyProTyrGlnLysGlySProValHisGluLysLysGluValLeu 99
         ||||||||||||||||||||||||||||||||||||||||||||||||
     384 GCCGGTCCTTACCAGAAAAAAGCTGTGCATGAAAAAAAGAAGTTTTG 430

seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2000.DAT:AAA15919

seq_documentation_block:
ID   AAA15919 standard; cDNA; 297 BP.
XX
AC   AAA15919;
XX
DT   12-JUN-2000 (first entry)
XX
DE   Human protein clone HP10357 coding sequence.
XX
KW   Human protein; hydrophobic domain; nutritional source; haematopoesis;
KW   cytokine production; cell proliferation; cell differentiation;
KW   immune deficiency; infectious disease; autoimmune disorder; asthma;
KW   multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
KW   allergic reaction; osteoporosis; osteoarthritis; periodontal disease;
KW   nervous system disorder; Alzheimer's disease; Parkinson's disease;
KW   Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;
KW   systemic cytokine damage; tissue differentiation; contraceptive; stroke;
KW   coagulation disorder; myocardial infarction; inflammatory condition;
KW   septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;
KW   nephritis; therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO20005367-A2.
XX
PD   03-FEB-2000.
XX
PF   22-JUL-1999;   99WO-JP03929.

PR   24-JUL-1998;   98JP-0208820.
PR   07-AUG-1998;   98JP-0224105.
PR   25-AUG-1998;   98JP-0238116.
PR   09-SEP-1998;   98JP-0254736.
PR   29-SEP-1998;   98JP-0275505.
XX
PA   (SAGA) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S, Kimura T;
XX
DR   WPI; 2000-182694/16.
DR   P-PSDB; AAY94861.
XX
PT   Novel human proteins having hydrophobic domains useful for treating
PT   osteoporosis, Alzheimer's disease, Parkinson's disease, asthma,
PT   multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke -
XX
PS   Claim 3; Page 217-218; 351pp; English.
XX
CC   This sequence encodes a human protein of the invention, which has
CC   hydrophobic domains. The protein of the invention. The DNA sequences can be used as a probe or as a
CC   genetic marker. The protein can also be used as a marker, and to identify
CC   potential genetic disorders. The DNA and protein can also be used as
CC   nutritional sources or supplements. The protein exhibits cytokine, cell
CC   proliferation, cell differentiation activities and induces production of
CC   other cytokines in certain cell populations. The protein also exhibits
CC   immune stimulating or immune suppressing activity. It can be used in the
CC   treatment of various immune deficiencies and disorders, and to treat
CC   infectious diseases caused by viral, bacterial, fungal or other
CC   infections. The protein is also used for treating autoimmune disorders
CC   such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
CC   arthritis. It is also useful in the treatment of allergic reactions and
CC   conditions such as asthma, and in immune suppression after organ
CC   transplantation. The protein is useful in regulation of haematopoiesis
CC   and consequently in the treatment of myeloid or lymphoid cell
CC   deficiencies. It is also used in compositions for tissue growth or
CC   regeneration. The protein is also used in the treatment of osteoporosis
CC   or osteoarthritis and in the treatment of periodontal disease and other
CC   tooth repair processes. The protein is used in the treatment of nervous
CC   system disorders such as Alzheimer's disease, Parkinson's disease, and
CC   Huntington's disease. They are useful for protection or regeneration and
CC   treatment of lung or liver fibrosis, reperfusion injury in various
CC   tissues, and conditions resulting from systemic cytokine damage. They are
CC   also used for promoting or inhibiting tissue differentiation. They are
CC   also used as contraceptives since they exhibit activin or inhibin related
CC   activities and as a fertility inducing therapeutic. They are used for
CC   treating various coagulation disorders and in treatment and prevention of
CC   conditions resulting from coagulation activities e.g. myocardial
CC   infarction or stroke. They also acts as receptors, receptor ligands or
CC   inhibitors or agonists of receptor/ligand interactions. They are used to
CC   treat inflammatory conditions such as septic shock, sepsis, ischaemia
CC   reperfusion injury, arthritis, and nephritis. They can be used to
CC   prevent tumours.
XX
SQ   Sequence 297 BP; 81 A; 68 C; 72 G; 76 T; 0 other;

alignment_scores:
           Quality: 504.00
           Ratio: 5.091
Percent Similarity: 100.000   Percent Identity: 98.990.

alignment_block:
US-09-801-115-2 x AAA15919

Align seg 1/1 to: AAA15919 from: 1 to: 297

       1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGT 50

      17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34

```
51  GAAAGGCCACGTGAAGAAGTGCTGCGGCTGGATATTATCAACTCACTGGTAA   100
    ||||||||||||||||||||||||||||||||||||||||||||||||||||

34  hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr     50
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
101 CAACAGTATTCATGCCTCATCGTAATCGTATCTGCTGGCCACTGAGAAAC    150

51  ThrThrLeuThrValGlyGlyIleGlyValPheAlaLeuValThrAlaValCy  67
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
151 ACAACATTGACACAGTTGGTGGGAGGGGTGTTTGCACTTGTGACAGCGTATG  200

67  sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuProPheAsnPros  84
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
201 CTGTCTTGCCGACGGGGCCCTTATTTACCGGAAGCTTCGTTCAATCCA  250

84  erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu        99
    |||||||||::|||||||||||||||||||||||||||||||||||||
251 GCGGTCCTTACCACGAAAGCCTGTGCATGAAAAAAAGAAGTTTG  297
```

---

alignment_scores:
```
            Quality:  504.00          Length:        99
            Ratio:      5.091          Gaps:           0
Percent Similarity: 100.000  Percent Identity:  98.990
```

Align seg 1/1 to: AAA15929   from: 1 to: 467

```
  1 MetAsnAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa  17
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
114 ATGGATAACGTGCAGCCGAAAAATAAAACATCGCCCCTTCTGTTCAGTGT  163

 17 lLysGlyHisValLysMetLeuArgLeuAspIleIleIleAsnSerLeuValT  34
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
164 GAAAGGCCACGTGAAGAAGTGCTGCGGCTGGATATTATCAACTCACTGGTAA  213

 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr     50
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
214 CAACAGTATTCATGCCTCATCGTAATCGTATCTGCTGGCCACTGAGAAAC  263

 51 ThrThrLeuThrValGlyGlyIleGlyValPheAlaLeuValThrAlaValCy  67
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
264 ACAACATTGACACAGTTGGTGGGAGGGGTGTTTGCACTTGTGACAGCGTATG  313

 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuProPheAsnPros  84
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
314 CTGTCTTGCCGACGGGGCCCTTATTTACCGGAAGCTTCGTTCAATCCA  363

 84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu        99
    |||||||||::|||||||||||||||||||||||||||||||||||||
364 GCGGTCCTTACCACGAAAGCCTGTGCATGAAAAAAAGAAGTTTG  410
```

XX
AC   AAX97826;
XX
DT   23-SEP-1999   (first entry)
XX
DE   Human secreted protein encoding cDNA #14.
XX
KW   Secreted protein; human; cytostatic; thrombotic; osteopathic; forensic;
KW   diagnostic; gene therapy; chromosome mapping; secretion vector; ss.
XX
OS   Homo sapiens.
XX
PN   WO9925825-A2.
XX
PD   27-MAY-1999.
XX
PF   13-NOV-1998;
XX
PR   04-SEP-1998;   98US-0099273.
PR   13-NOV-1997;   97US-0066677.
PR   17-DEC-1997;   97US-0069957.
PR   09-FEB-1998;   98US-0074121.
PR   13-APR-1998;   98US-0081563.
PR   10-AUG-1998;   98US-0096116.
XX
PA   (GEST) GENSET.
XX
PI   Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR   WPI; 1999-347472/29.
DR   P-PSDB; AAY36142.
XX
PT   Extended cDNAs encoding secreted proteins
XX
PS   Example 28; Page 173-174; 307pp; English.
XX
CC   AAX97813-X97906 represent extended cDNA's which encode novel human
CC   secreted proteins (see AAY36129-Y36222) and which have cytostatic,
CC   thrombotic and osteopathic activity. The extended cDNAs can be used to
CC   express secreted proteins or parts of them or to obtain antibodies
CC   capable of binding to the secreted proteins. They may also be used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   Uses also include design of expression vectors and secretion vectors.
XX
SQ   Sequence 500 BP; 135 A; 102 C; 125 G; 129 T; 9 other;

alignment_scores:
       Quality: 489.00   Length: 100
       Ratio: 4.990   Gaps: 1
Percent Similarity: 98.000   Percent Identity: 98.000

alignment_block:
US-09-801-115-2 x AAX97826

Align seq 1/1   to: AAX97826   ..
Align seq 1/1   to: AAX97826   from: 1 to: 500

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa  17
    |||||||||||||||||||||||||||||||||||||||||||||||||
132 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGT  181

 17 LLysMetLeuArgLeuAspIleIleAsnSerLeuVal  33
    |||||||||||||||||||||||||||||||||||||
182 GAAAGGCCACGTGAYAGATGCTCCGGCTGGATATTATCAACTCACTGGTA  231

 34 ThrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluTh  50
    |||||||||||||||||||||||||||||||||||||||||||||||||
232 ACAACAGTATTCATCGTCATCGTATCTGTGTTGGCACTGATACCAGAAAC  281

 50 rThrThrLeuThrValGlyValPheAlaLeuValThrAlaValC  67
    ||||||||||||||||||||||||||||||||||||||||||||
282 CACAACAVTGACACAGTTGGTGGAGGGGTGTTTGCACTGTGTGACAGCAGTAT  331

 67 ysCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnPro  83
    |||||||||||||||||||||||||||||||||||||||||||||||||
332 GCTGTCTTGCCGACGGGCCCCTTATTTACCGGGAGCTTCTGTTCAAATCC  381

 84 SerGlyProTyrGlnLysHisProValHisGluLysSerGluValLeu  99
    ||||||||||||||||||||||||||||||||||||||||||||||||
382 AGCGGCGTCCTTACCAGAAAAAAGCCTGTGCATGAAAAAAGAAGTTTTG  429
```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAA38007

seq_documentation_block:
ID   AAA38007 standard; cDNA; 459 BP.
XX
AC   AAA38007;
XX
DT   22-AUG-2000   (first entry)
XX
DE   UCK-2 nucleotide sequence.
XX
KW   UCK-2; chemotaxis factor; immunocyte; haemopoiesis stimulant; tumour;
KW   radiotherapy; chemotherapy; human; ss.
XX
OS   Homo sapiens.
XX
PN   CN1244584-A.
XX
PD   16-FEB-2000.
XX
PF   14-MAY-1999;   99CN-0107284.
XX
PR   14-MAY-1999;   99CN-0107284.
XX
PA   (UYBE-) UNIV BEIJING MEDICAL.
XX
PI   Ma D, Han W, Zhang Y;
XX
DR   WPI; 2000-388170/34.
DR   P-PSDB; AAY98143.
XX
PT   Chemotaxic factor useful for treatment and diagnosis of immunocyte
PT   disorders - has immunocyte chemotaxic stimulating factor
XX
PS   Example 4; Fig 2; 31pp; Chinese.
XX
CC   This sequence represents an UCK-2 cDNA sequence encoding a chemotaxis
CC   factor polypeptide. The UCK-2 protein exhibits immunocyte chemotaxis
CC   activity and a haemopoiesis stimulating effect. The invention relates to
CC   UCK proteins, their encoding nucleotide sequences and antibodies and
CC   antagonists against the proteins. The nucleotide and protein sequences
CC   are useful for the preparation of a composition for the diagnosis and
CC   treatment of diseases associated with abnormal immunocyte function and
CC   low haemopoiesis function caused by radiotherapy and chemotherapy used to
CC   treat tumours and other diseases.
XX
SQ   Sequence 459 BP; 123 A; 99 C; 96 G; 141 T; 0 other;

alignment_scores:
       Quality: 471.50   Length: 152
       Ratio: 4.763   Gaps: 1
Percent Similarity: 65.132   Percent Identity: 65.132

alignment_block:
US-09-801-115-2 x AAA38007

Align seq 1/1   to: AAA38007   from: 1 to: 459

```
 1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa  17
   |||||||||||||||||||||||||||||||||||||||||||||||||
 1 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGT  50

17 LLysGlyHisValValLysMetLeuArg..........................  25
   |||||||||||||||||||||||||
```

```
PT  Proteins, polynucleotides, vectors, host cells and antibodies used to
PT  diagnose, treat or prevent immune, reproductive, smooth muscle,
PT  neurological, gastrointestinal, developmental and cell proliferative
PT  disorders -
XX
PS  Claim 9; Page 210; 229pp; English.
XX
CC  AA256698 to AA256776 encode AAY57877 to AAY57955 which represent human
CC  transmembrane proteins designated HTMPN-1 to HTMPN-79, respectively.
CC  The transmembrane proteins have immunospecific, antiproliferative and
CC  neuroprotective activities. The human transmembrane proteins,
CC  polynucleotides encoding them and other compositions and methods from
CC  the present invention, can be used for the diagnosis, treatment or
CC  prevention of immune, reproductive, smooth muscle, neurological,
CC  gastrointestinal, developmental and cell proliferative disorders. The
CC  HTMPN's can be used to treat or prevent disorders associated with a
CC  decreased expression or activity of HTMPN.
XX
SQ  Sequence 637 BP; 169 A; 134 C; 143 G; 191 T; 0 other;
```

```
alignment_scores:
         Quality: 471.50            Length:     152
         Ratio:   4.763             Gaps:       1
 Percent Similarity: 65.132    Percent Identity: 65.132

alignment_block:
US-09-801-115-2 x AA256747

Align seg 1/1  to: AA256747  from: 1  to: 637

  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa  17
    |||||||||||||||||||||||||||||||||||||||||||||||||||
111 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGT   160

 17 LysGlyHisValLysMetLeuArg.......................  25
    |||||||||||||||||||
161 GAAAGGCCACGTGAAGAATGCTCGGGCTGGCACTAACTGTGACATCTATGA  210

 25 ..................................................  25
211 CCTTTTTATCATCGCACAAGCCCCTGAACCAATATATTGTTATCACTGGA   260

 25 ..................................................  25
261 TTTGAAGTCACCGTTATCTTATTATTTTCATACTTTTATATGTGACTCAGACT 310

 26 ...............LeuAspIleIleAsnS  31
                  ||||||||||||||
311 TGATCGATTAAATGAAGTGGTTATTTTGCCTTTGCTTGATATTATCAACT   360

 31 erLeuValThrThrValGlyGlyValPheMetLeuIleValSerValLeuAlaLeuIle 47
    |||||||||||||||||||||||||||||||||||||||||||||||||
361 CACTGGTAACAACACAGTATTCATCGCTCATCGTATGTTGCTTGGCACTGATA 410

 48 ProGluThrThrLeuThrValGlyGlyValPheAlaLeuValTh  64
    ||||||||||||||||||||||||||||||||||||||||||
411 CCAGAAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGAC   460

 64 rAlaValCysCysLeuAlaAspGlyAlaLeuLeuIleTyrArgLysLeuLeuP 81
    |||||||||||||||||||||||||||||||||||||||||||||||||
461 AGCAGTATGCTGTCTGCCGACGGGGCCCTTATTTACCGGAAGCTTCTGT    510

 81 heAsnProSerGlyProTyrGlnLysLysProValHisGluLysLysGlu  97
    ||||||||||||||||||||||||||||||||||||||||||||||||
511 TCAATCCCAGCGGGTCCTTACCAGAAAAAGCCTGTGCATGAAAAAAAGAA   560

 98 ValLeu  99
    ||||||
561 GTTTTG  566
```

```
seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AA256747

seq_documentation_block:
ID  AA256747 standard; cDNA; 637 BP.
XX
AC  AA256747;
XX
DT  23-MAR-2000  (first entry)
XX
DE  Human transmembrane protein HTMPN-50 encoding cDNA.
XX
KW  Human; transmembrane protein; HTMPN; diagnosis; immunospecific;
KW  antiproliferative; neuroprotective; immune disorder;
KW  reproductive disorder; smooth muscle disorder; neurological disorder;
KW  gastrointestinal disorder; developmental disorder; ss.
OS  Homo sapiens.
XX
PN  WO9961471-A2.
XX
PD  02-DEC-1999.
XX
PF  28-MAY-1999;    99WO-US11904.
XX
PR  29-MAY-1998;    98US-0087260.
PR  02-JUL-1998;    98US-0091674.
PR  02-OCT-1998;    98US-0102954.
PR  24-NOV-1998;    98US-0109869.
XX
PA  (INCY-) INCYTE PHARM INC.
XX
PI  Tang YT,  Lal P,  Hillman JL,  Yue H,  Guegler KJ,  Corley NC;
PI  Bandman O,  Patterson C,  Gorgone GA,  Kaser MR,  Baughn MR;
PI  Au-Young J;
XX
DR  WPI; 2000-072605/06.
DR  P-PSDB; AAY57926.
XX
```

```
 51 GAAAGGCCACGTGAAGAATGCTCGGGCTGGCACTAACTGTGACATCTATGA  100
 25 ..................................................   25

101 CCTTTTTATCATCGCACAAGCCCCTGAACCAATATATTGTTATCACTGGA   150
 25 ..................................................   25

151 TTTGAAGTCACCGTTATCTTATTATTTTCATACTTTTATATGTGACTCAGACT 200
 26 ...............LeuAspIleIleAsnS  31

201 TGATCGATTAAATGAAGTGGTTATTTTGCCTTTGCTTGATATTATCAACT   250
 31 erLeuValThrThrValGlyGlyValPheMetLeuIleValSerValLeuAlaLeuIle 47

251 CACTGGTAACAACACAGTATTCATCGCTCATCGTATGTTGCTTGGCACTGATA 300
 48 ProGluThrThrLeuThrValGlyGlyValPheAlaLeuValTh  64

301 CCAGAAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGAC   350
 64 rAlaValCysCysLeuAlaAspGlyAlaLeuLeuIleTyrArgLysLeuLeuP 81

351 AGCAGTATGCTGTCTGCCGACGGGGCCCTTATTTACCGGAAGCTTCTGT    400
 81 heAsnProSerGlyProTyrGlnLysLysProValHisGluLysLysGlu  97

401 TCAATCCCAGCGGGTCCTTACCAGAAAAAGCCTGTGCATGAAAAAAAGAA   450
 98 ValLeu  99

451 GTTTTG  456
```

OM of: US-09-801-115-2 to: GenEmbl:* out_format : pfs

Q:~/cgn2.1/USPTO_spool/US09801115/runat.19072002.182249.6115/app_query.fasta_1.156
Date: Jul 19, 2002 7:57 PM
About: Results were produced by the Gencore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL-frame+_p2n.model: -DEV-xlh
-Q=~/cgn2.1/USPTO_spool/US09801115/runat.19072002.182249.6115/app_query.fasta_1.156
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=1.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human4o.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000000 -USER-US09801115_@CGN1_1.4008
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-801-115-2
Query length: 99
Database: GenEmbl:*
Database sequences: 1797656
Database length: 1873333701
Search time (sec): 1842.490000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation
gb_pat:AX061624	+	508.00	1154.28	6.2e-56	515	! AX061624 Sequence 29 from Paten
gb_pat:AX061624	+	508.00	1154.01	6.5e-56	515	! AX061624 Sequence 29 from Paten
gb_pr:AF099895	+	508.00	1154.01	6.5e-56	530	! AF099895 Homo sapiens chemokine
gb_vi:AF057306	+	471.50	1068.31	3.9e-51	655	! AF057306 Homo sapiens transmemb
gb_pat:AX061665	+	471.50	1067.93	4.0e-51	669	! AX061665 Sequence 70 from Paten
gb_pr:BC004380	+	471.50	1067.67	4.2e-51	688	! BC004380 Homo sapiens, clone MG
gb_pat:BC004380	+	471.50	1067.67	4.2e-51	689	! AF135380 Homo sapiens chemokine
gb_sts:G30204	+	366.00	828.43	8.8e-38	485	! G30204 human STS SHGC-36487, se
gb_pr:AF145216	+	339.50	765.63	2.8e-34	593	! AF145216 Homo sapiens chemokine
gb_pr:AF135381	+	331.00	749.02	2.3e-33	434	! AF135381 Homo sapiens chemokine
gb_pr:AX079435	+	318.50	719.05	1.1e-31	495	! AX079435 Sequence 179 from Pate
gb_ro:AF253064	+	275.00	618.55	4.3e-26	523	! AF253064 Rattus norvegicus chem
gb_pat:AX330610	-	245.00	558.27	9.9e-23	207	! AX330610 Sequence 1119 from Pat
gb_pat:AX330787	-	245.00	558.27	9.9e-23	207	! AX330787 Sequence 1296 from Pat
gb_pat:AX331008	-	245.00	558.27	9.9e-23	207	! AX331008 Sequence 1517 from Pat
gb_ro:AF253065	+	238.50	532.16	2.8e-21	682	! AF253065 Rattus norvegicus chem
gb_ro:AY047360	+	227.50	508.59	5.8e-20	568	! AY047360 Mus musculus chemokine
gb_htg:AC018589	+	215.00	427.51	1.9e-15	151620	! AC018589 Homo sapiens clone
gb_pr:AC010542	+	215.00	425.47	2.5e-15	188460	! AC010542 Homo sapiens chromo
gb_htg:AC018557	+	215.00	425.40	2.5e-15	189821	! AC018557 Homo sapiens chromo
gb_htg:AC010289	+	142.00	263.88	2.5e-06	97083	! AC010289 Homo sapiens chromos
gb_ro:AY045597	+	139.50	307.00	9.0e-09	527	! AY045597 Mus musculus chemokine
gb_pat:AX252052	-	108.00	211.38	0.0021	792	! AX252052 Sequence 313 from Pat
gb_pat:AX344430	+	108.00	211.38	0.0021	6283	! AX344430 Sequence 277 from Pat
gb_pat:AX348835	-	108.00	211.38	0.0021	6283	! AX348835 Sequence 293 from Pat
gb_pat:AX252053	-	102.00	197.58	0.0121	6283	! AX252053 Sequence 314 from Pat
gb_pat:AX344431	+	102.00	197.58	0.0121	6283	! AX344431 Sequence 278 from Patt
gb_pat:AX348836	-	102.00	197.58	0.0121	6283	! AX348836 Sequence 294 from Pat
gb_htg:AC000090	+	88.00	146.87	8.10	45416	! AC000090 Homo sapiens Chromos
gb_pr:AC010542	+	86.50	182.52	0.0838	700	! AF278577 Homo sapiens chemokine
gb_ba:AE004148	-	85.50	152.84	3.77	13007	! AE004148 Vibrio cholerae chro
gb_vi:AF105732	-	83.50	174.46	0.2355	792	! AF105732 HIV-1 isolate A-RII-07
gb_vi:AF105741	-	83.50	174.46	0.2355	792	! AF105741 HIV-1 isolate A-RII-40
gb_pat:AR166758	+	82.50	137.06	28.52	37948	! AR166758 Sequence 11 from pat
gb_ro:AF401530	-	81.00	176.32	0.1855	508	! AF401530 Mus musculus chemokin
gb_vi:AF094041	-	81.00	170.90	0.3716	627	! AF094041 HIV-1 isolate 313-9-3-
gb_pr:AL591368	-	79.50	124.06	151.18	64425	! AL591368 Human DNA sequence f
gb_pr:AC023268	-	79.50	116.49	398.91	144457	! AC023268 Homo sapiens chromo
gb_htg:AC016329	-	79.50	116.11	418.81	150430	! AC016329 Homo sapiens clone
gb_htg:AC016329	-	79.50	116.11	418.81	150430	! AC016329 Homo sapiens clone
gb_htg:AL355380	+	79.50	109.17	1.0e+03	315681	! AL355380 Homo sapiens chromo

seq_name: gb_pat:AX061624

seq_documentation_block:
LOCUS   AX061624   515 bp   DNA   linear   PAT 24-JAN-2001
DEFINITION Sequence 29 from Patent WO0100806.
ACCESSION   AX061624
VERSION   AX061624.1   GI:12406709
KEYWORDS   .
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 515)
  AUTHORS   dumas milne Edwards,J.B., Bougueleret,L. and Jobert,S.
  TITLE   Complementary dna's encoding proteins with signal peptides
  JOURNAL   Patent: WO 0100806-A 29 04-JAN-2001;
            GENSET (FR)
FEATURES   Location/Qualifiers
     source   1..515
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     sig_peptide   144..287
                     /note="von Heijne matrix score 4.10 seq VFMLIVSLALIP/ET"
     CDS   144..443
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC24985.1"
                     /db_xref="GI:12406710"
                     /translation="MDNVQPKIKHRPFCFSVKGHVKMLRLDIINSLVTTVFMLIVSVL
                     ALIPETTLIPVGGGVFALVYPAVCCLADGALIYRKLLFNPSGPYQKKPVHEKKEVL"
     polyA_signal   457..462
     polyA_site   500..515
BASE COUNT   143 a   106 c   135 g   130 t   1 others
ORIGIN

US-09-801-115-2 x AX061624

alignment_block:
US-09-801-115-2 to: AX061624 from: 1 to: 515

alignment_scores:
   Quality: 508.00   Length: 99
   Ratio: 5.131   Gaps: 0
   Percent Similarity: 100.000   Percent Identity: 100.000

Align seg 1/1 to: AX061624 from: 1 to: 515

      1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    144 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGTTTCAGTGT 193

     17 LLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    194 GAAAGGCCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 243

     34 hrThrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGlu 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    244 CAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACC 293

     51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    294 ACAACATTGACACAGTTGGTGGAGGGGGTGTTTGCACTGGTGACAGCAGT 343

     67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnPros 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    344 CTGTGTTGCCGACGGGGCGCCCTTATTTACCGGAAGCTTCTGTTCAATCCA 393

     84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
        |||||||||||||||||||||||||||||||||||||||
    84

```
           |||||||||||||||||||||||||||||||||||||||||||||
       394 GCGGTGCTTACCAGAAAAAAGCCTGTGCATGAGAAAAAGAAGTTTTG 440
```

seq_name: gb_pr:AF096895

seq_documentation_block:
```
LOCUS       AF096895                 530 bp    mRNA    linear   PRI 18-JUL-2001
DEFINITION  Homo sapiens chemokine-like factor 1 (CKLF1) mRNA, complete cds.
ACCESSION   AF096895
VERSION     AF096895.2  GI:9989691
KEYWORDS    FLI_cDNA.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   Han,W., Lou,Y., Tang,J., Zhang,Y., Chen,Y., Li,Y., Gu,W., Huang,J.,
            Gui,L., Tang,Y., Li,F., Song,Q., Di,C., Wang,L., Shi,Q., Sun,R.,
            Xia,D., Rui,M., Tang,J. and Ma,D.
  TITLE     Molecular cloning and characterization of chemokine-like factor 1
            (CKLF1), a novel human cytokine with unique structure and potential
            chemotactic activity
  JOURNAL   Biochem J. 357 (Pt 1), 127-135 (2001)
   MEDLINE   21308461
    PUBMED   11415443
REFERENCE   2  (bases 1 to 530)
  AUTHORS   Han,W.L., Li,Y., Zhang,Y.M., Di,C.H., Song,Q.S. and Ma,D.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-1998) Immunology, Beijing Medical University, 38
            Xue Yuan Rd., Beijing 100083, China
REFERENCE   3  (bases 1 to 530)
  AUTHORS   Han,W.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2000) Immunology, Peking University Health
            Science Center, Xueyuan Road 38, Beijing 100083, China
  REMARK    Sequence update by submitter
COMMENT     On Sep 7, 2000 this sequence version replaced gi:6288733.
FEATURES             Location/Qualifiers
     source          1..530
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="U937"
                     /note="exposed to phytohaemagglutinin (PHA)"
     gene            1..530
                     /gene="CKLF1"
                     /note="UCK-1"
     CDS             148..447
                     /gene="CKLF1"
                     /note="increased expression in PHA stimulated cells;
                     expression inhibited by IL-10"
                     /codon_start=1
                     /product="chemokine-like factor 1"
                     /protein_id="AAF06722.1"
                     /db_xref="GI:6288734"
                     /translation="MDNVQPKIKHRPFCFSVKGHVKMLRLDIINSLVTTVFMLIVSVL
                     ALIPETTLTVGGGVFALVTAVCCLADGALIYRKLLFNPSGPYQKKPVHEKKEVL"
BASE COUNT      156 a    108 c    137 g    129 t
ORIGIN
```
```
         1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
           |||||||||||||||||||||||||||||||||||||||||||||||||||
       148 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGT 197
```

alignment_scores:
```
Quality:  508.00              Length:  99
Ratio:    5.131               Gaps:    0
Percent Similarity: 100.000   Percent Identity: 100.000
```
alignment_block:
US-09-801-115-2 x AF096895

Align seg 1/1 to: AF096895 from: 1 to: 530

```
        17 LLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
           |||||||||||||||||||||||||||||||||||||||||||||||||||
       198 GAAAGGGCACGTGAAGATGCTGCGGCTGGATATTATCAACTCACTGGTAA 247
```
```
        34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
           |||||||||||||||||||||||||||||||||||||||||||||||||||
       248 CAACAGTATTCATCGTCAGTATCGTTGGTGTTGGGCACTGATACCAGAGACC 297
```
```
        51 ThrThrLeuThrValGlyGlyGlyValPheAlaLeuValThrAlaValCy 67
           |||||||||||||||||||||||||||||||||||||||||||||||||||
       298 ACAACAATTGACAGTTTGGTGGGAGGGGTGTTTGCACTTGTGACACGAATG 347
```
```
        67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnProS 84
           |||||||||||||||||||||||||||||||||||||||||||||||||||
       348 CTGTCTTGCCGACGGGGGCCTTTAATTTACCGGAAGCTTCTGTTCAATCCA 397
```
```
        84 erGlyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
           ||||||||||||||||||||||||||||||||||||||||||||||
       398 GCGGGTGCTTACCAGAAAAAGCCTGTGCATGAGAAAAAGAAGTTTTG 444
```

seq_name: gb_pr:AF057306

seq_documentation_block:
```
LOCUS       AF057306                 655 bp    mRNA    linear   PRI 31-DEC-1999
DEFINITION  Homo sapiens transmembrane proteolipid (C32) mRNA, complete cds.
ACCESSION   AF057306
VERSION     AF057306.1  GI:6648618
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Zhang,J.S., Nelson,M., Wang,L. and Smith D.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-APR-1998) Pathology and Lab. Medicine, Mayo Clinic
            and Foundation, Rochester, MN 55905, USA
FEATURES             Location/Qualifiers
     source          1..655
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="AsPC-1"
                     /tissue_type="pancreas"
                     /note="adenocarcinoma"
     gene            1..655
                     /gene="C32"
     CDS             132..590
                     /gene="C32"
                     /note="down regulated upon cell differentiation induced by
                     sodium butyrate"
                     /codon_start=1
                     /product="transmembrane proteolipid"
                     /protein_id="AAF21255.1"
                     /db_xref="GI:6648619"
                     /translation="MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEP
                     YIVIGFEVTVILFILLYILRLDRLMKWLFWPLLDIIINSLVTTVFMLIVSVLALIPE
                     TTTLIVGGGVFALVTAVCCLADGALIYRKLLFNPSGPYQKKPVHEKKEVL"
BASE COUNT      171 a    137 c    153 g    194 t
ORIGIN
```

alignment_scores:
```
Quality:  471.50              Length:  152
Ratio:    4.763               Gaps:    1
Percent Similarity: 65.132    Percent Identity: 65.132
```
alignment_block:
US-09-801-115-2 x AF057306

Align seg 1/1 to: AF057306 from: 1 to: 655

```
         1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
           |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
132 ATGGATAACGTGCAGCCGAAAAATAAAACATCGCCCCTTCTGCTTCAGTGT 181
    ||||||||||||||||||||||||||||||||||||||||||||||||||
    lLysGlyHisValLysMetLeuArg                          25

17  lLysGlyHisValLysMetLeuArg                           25
182 GAAAGGCCACGTGAAGATCGCGGCTGCCACTAACTGTGACATCTATGA   231
    ||||||||||||||||||||||||||||||||||||||||||||||||
                                                         25
232 CCTTTTTTTATCATCGCACAAGCCCCTGAACCATATATTGTTATCACTGGA 281
    ||||||||||||||||||||||||||||||||||||||||||||||||||
25  ...................................................  25
282 TTTGAAGTCACCGTTATCTTTATTTTTCATACTTTTATATGTACTCAGACT 331
    ||||||||||||||||||||||||||||||||||||||||||||||||||
25  ...................................................  25
332 TGATCGATTAATGAAGTGGTTATTTTTGGGCCTTTGCTTGATATTATCAACT 381
    ||||||||||||||||||||||||||||||||||||||||||||||||||
    LeuAspIleIleAsnS                                     31
31  erLeuValThrThrValPheMetLeuIleIleValSerValLeuAlaLeuIle 47
382 CACTGGTAACAACAGTATTCATCGTCATCGTGTTGTTGGCCACTGATA     431
    ||||||||||||||||||||||||||||||||||||||||||||||
48  ProGluThrThrThrLeuThrHisValGlyLeuGlyLeuGlyValAlaLeuValThr 64
432 CCAGAAACCACAACATTGACAGTTGGGTGGGAGGGGTGTTTGCACTTGTGAC 481
    ||||||||||||||||||||||||||||||||||||||||||||||||||
64  rAlaValCysCysLeuAlaAspGlyAlaLeuIleLeuTyrArgLysLeuLeuP 81
482 AGCAGTATCCTGTCTTGCGGACGGGGCCCTTATTTACCGGAAGCTTCTGT 531
    ||||||||||||||||||||||||||||||||||||||||||||||||||
81  heAsnProSerGlyProTyrGlnLysLysProValHisGluLysLysGlu 97
532 TCAATCCAGCGGGTCCTTACCAGAAAAAGCCCTGTGCATGAAAAAAAGAA 581
    ||||||||||||||||||||||||||||||||||||||||||||||||||
98  ValLeu 99
582 GTTTTG 587
    ||||||
```

seq_name: gb_pat:AX061665

```
LOCUS       AX061665    669 bp   DNA   linear   PAT 24-JAN-2001
DEFINITION  Sequence 70 from Patent WO0100806.
ACCESSION   AX061665
VERSION     AX061665.1  GI:1240G789
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 669)
AUTHORS     dumas mine Edwards,J.B., Bouqueleret,L. and Jobert,S.
TITLE       Complementary dna's encoding proteins with signal peptides
JOURNAL     Patent: WO 0100806-A 70 04-JAN-2001;
            GENSET (FR)
FEATURES             Location/Qualifiers
     source          1..669
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             140..598
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25025.1"
                     /db_xref="GI:12406790"
                     /translation="MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEP
                     YIVITGPEVTVILFFILLYVLRDLRMKWLFWPLLDIINSLVTTVFMLIVSVLAIPE
                     TTFLIVVGGGVFALVTRAVCCLADGALIYRKLLFNPSGPYQKKPVHEKKEVL"
     sig_peptide     140..442
                     /note="Von Heijne matrix score 4.10 seq VFMLIVSVLALIP/ET"
     polyA_signal    630..635
     polyA_site      655..669
BASE COUNT  178 a   139 c   158 g   194 t
```

ORIGIN

```
alignment_scores:
           Quality:   471.50           Length:    152
           Ratio:     4.763            Gaps:        1
           Percent Similarity: 65.132   Percent Identity: 65.132

alignment_block:
US-09-801-115-2 x AX061665

Align seq 1/1   to: AX061665   from: 1  to: 669
US-09-801-115-2  to: AX061665  ..

1   MetAspAsnValGlnProLysIleLeuHisArgProPheCysPheSerVa 17
140 ATGGATAACGTGCAGCCGAAAAATAAAACATCGCCCCTTCTGCTTCAGTGT 189
    ||||||||||||||||||||||||||||||||||||||||||||||||||
17  lLysGlyHisValLysMetLeuArg.......................... 25
190 GAAAGGCCACGTGAAGATCGCGGCTGCCACTAACTGTGACATCTATGA .. 239
    ||||||||||||||||||||||||||||||||||||||||||||||||
25  .................................................. 25
240 CCTTTTTTTATCATCGCACAAGCCCCTGAACCATATATTGTTATCACTGGA 289
    ||||||||||||||||||||||||||||||||||||||||||||||||||
25  .................................................. 25
290 TTTGAAGTCACCGTTATCTTTATTTTTCATACTTTTATATGTACTCAGACT 339
    ||||||||||||||||||||||||||||||||||||||||||||||||||
25  .................................................. 25
340 TGATCGATTAATGAAGTGGTTATTTTTGGGCCTTTGCTTGATATTATCAACT 389
    ||||||||||||||||||||||||||||||||||||||||||||||||||
26  .................LeuAspIleIleAsnS                    31
31  erLeuValThrThrValPheMetLeuIleIleValSerValLeuAlaLeuIle 47
390 CACTGGTAACAACAGTATTCATCGTCATCGTGTTGTTGGCCACTGATA 439
    ||||||||||||||||||||||||||||||||||||||||||||||
48  ProGluThrThrThrLeuThrHisValGlyLeuGlyLeuGlyValAlaLeuValThr 64
440 CCAGAAACCACAACATTGACAGTTGGGTGGGAGGGGTGTTTGCACTTGTGAC 489
    ||||||||||||||||||||||||||||||||||||||||||||||||||
64  rAlaValCysCysLeuAlaAspGlyAlaLeuIleLeuTyrArgLysLeuLeuP 81
490 AGCAGTATCCTGTCTTGCGGACGGGGCCCTTATTTACCGGAAGCTTCTGT 539
    ||||||||||||||||||||||||||||||||||||||||||||||||||
81  heAsnProSerGlyProTyrGlnLysLysProValHisGluLysLysGlu 97
540 TCAATCCAGCGGGTCCTTACCAGAAAAAGCCCTGTGCATGAAAAAAAGAA 589
    ||||||||||||||||||||||||||||||||||||||||||||||||||
98  ValLeu 99
590 GTTTTG 595
    ||||||
```

seq_name: gb_pr:BC004380

```
LOCUS       BC004380    688 bp   mRNA   linear   PRI 12-JUL-2001
DEFINITION  Homo sapiens, clone MGC:10658 IMAGE:3639550, mRNA, complete cds.
ACCESSION   BC004380
VERSION     BC004380.1  GI:13325133
KEYWORDS    MGC.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 688)
AUTHORS     Strausberg,R.
TITLE       Direct Submission
JOURNAL     Submitted (12-MAR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
```

COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution Information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 13 Row: f Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 9989692.
FEATURES   Location/Qualifiers
     source   1..688
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MGC:10658 IMAGE:3639550"
                     /tissue_type="Pancreas, adenocarcinoma"
                     /clone_lib="NIH_MGC_39"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS   137..595
                     /codon_start=1
                     /product="Unknown (protein for MGC:10658)"
                     /protein_id="AAH04380.1"
                     /db_xref="GI:13325134"
                     /translation="MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEP
                     YIVITGFEVTVILFFILLYVLRLDRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPE
                     TTTLTVGGGVFALVTAVCCLADGALIYRKLLFNPSGPYQKKPVHEKKEVL"
BASE COUNT   201 a   138 c   154 g   195 t
ORIGIN

alignment_scores:
   Quality: 471.50   Length: 152
   Ratio: 4.763   Gaps: 1
   Percent Similarity: 65.132   Percent Identity: 65.132

alignment_block:
us-09-801-115-2 x BC004380

Align seg 1/1   to: BC004380   from: 1   to: 688

1   MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa   17
    |||||||||||||||||||||||||||||||||||||||||||||||||
137 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGT 186

17 lLysGlyHisValLysMetLeuArg.....................   25
    ||||||||||||||||||||||||||
187 GAAAGGCCACGTGAAGATGCTGCGCGTGCACACTCATGA.......... 236

25   25
237 CCTTTTTATCATCGCACAAGCCCCTGAACCAATATATTGTTATCACTGGA 286

25   25
287 TTTGAAGTCACCGTTATCTTATTTTTCATTACTTTTATATGTACTCAGACT 336

26 ....LeuAspIleIleAsn   31
                                 ||||||||||||||
337 TGATCGATTAATGAAGTGGTTATTTTTGGCCTTTGCCTTGATATTAATCAACT 386

---

seq_name: gb_pr:AF135380

LOCUS   AF135380   689 bp   mRNA   linear   PRI 07-SEP-2000
DEFINITION Homo sapiens chemokine-like factor 2 (CKLF2) mRNA, complete cds,
            alternatively spliced.
ACCESSION   AF135380
VERSION   AF135380.2 GI:9989692
KEYWORDS   .
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 689)
  AUTHORS   Han,W.L., Gu,W., Li,Y., Zhang,Y., Song,Q., Di,C. and Ma,D.
  TITLE   Direct Submission
  JOURNAL   Submitted (17-MAR-1999) Immunology, Beijing Medical University, Xue
            Yuan Road, Beijing 100083, China
REFERENCE   2 (bases 1 to 689)
  AUTHORS   Han,W.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-SEP-2000) Immunology, Peking University Health
            Science Center, Xueyuan Road 38, Beijing 100083, China
  REMARK   Sequence update by submitter
COMMENT   On Sep 7, 2000 this sequence version replaced gi:6630853.
FEATURES   Location/Qualifiers
     source   1..689
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="U937"
                     /note="from PHA stimulated cells"
     gene   1..689
                     /gene="CKLF2"
                     /note="UCK"
     gene   148..606
                     /gene="CKLF2"
                     /note="UCK-2; alternatively spliced"
     CDS
                     /gene="CKLF2"
                     /note="UCK-2"
                     /codon_start=1
                     /product="chemokine-like factor 2"
                     /protein_id="AAF19599.1"
                     /db_xref="GI:6630854"
                     /translation="MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEP
                     YIVITGFEVTVILFFILLYVLRLDRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPE
                     TTTLTVGGGVFALVTAVCCLADGALIYRKLLFNPSGPYQKKPVHEKKEVL"
BASE COUNT   195 a   140 c   161 g   193 t
ORIGIN

31 erLeuValThrThrValPheMetLeuIleValSerValLeuAlaLeuIle   47
    ||||||||||||||||||||||||||||||||||||||||||||||||||
387 CACTGGTAACAACACAGTATTCATGCTATCAGTGCTTCATCGTATCGTATCTGCCGTGATA 436

48 ProGluThrThrThrLeuThrValGlyGlyGlyValPheAlaLeuValTh   64
    ||||||||||||||||||||||||||||||||||||||||||||||||||
437 CCAGAAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGAC 486

64 rAlaValCysCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuP   81
    ||||||||||||||||||||||||||||||||||||||||||||||||||
487 AGCAGTATGCTGTGTTGCCGACCGGCGCCCTTAATTACGGAAGCCTTCGT 536

81 heAsnProSerGlyProTyrGlnLysLysProValHisGluLysLysGlu   97
    ||||||||||||||||||||||||||||||||||||||||||||||||||
537 TCAATCCCAGCGGTTCCTTACCAGAAAAAGCCTGTCATGAAAAAAAAGAA 586

98 ValLeu   99
    ||||||
587 GTTTTG 592

alignment_scores:
   Quality: 471.50   Length: 152
   Ratio: 4.763   Gaps: 1
   Percent Similarity: 65.132   Percent Identity: 65.132

alignment_block:
US-09-801-115-2 x AF135380

Align seg 1/1 to: AF135380 from: 1 to: 689

```
  1 MetAsnAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
148 ATGGATAACGTGCAGCCCGGAAAATAAAACATCGCCCTTCTGCTTCAGTGT 197

 17 LysGlyHisValLysMetLeuArg.......................... 25
    |||||||||||||||||||||||
198 GAAAGGCCACGTGAAGATGCTGCGGTCGCCACTAAACTGCAGATCTATGA 247

 25 .................................................. 25
248 CCTTTTTTATCATGCGACAAAGCCCTGAACCAATATTATTGTTATCACTGGA 297

 25 .................................................. 25
298 TTTGAAGTCACCGTTATCTTATTTTTTCATTACTTTTATATAGTACTCAGACT 347

 26 .............................LeuAspIleIleAsnS 31
                                 ||||||||||||||||
348 TGATCGATTAATGAAGTGGTTATTTTGCCTTTTGCTTGATATTATCAACT 397

 31 erLeuValThrThrValPheMetLeuIleValSerValLeuAlaLeuIle 47
398 CACTGGTAACAACAGTAATTCATCGTACTCATCGTAATCGTGTTGGCACTGATA 447

 48 ProGluThrThrLeuThrHisValGlyGlyGlyValPheAlaLeuValTh 64
448 CCAGAAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGAC 497

 64 rAlaLeuValCysCysLeuAlaAspGlyAlaLeuIleThrArgLysLeuLeuP 81
498 AGCAGTATGCTGTCTTGCCGACGGGCGCCTTATTTACCGGAAGCTTCTGT 547

 81 heAsnProSerGlyProTyrGlnLysLysProValHisGluLysLysGlu 97
548 TCAATCCCAGGCGGTCCTTACCAGAAAAAGCCTGTGCATGAAAAAAAAGAA 597

 98 ValLeu 99
598 GTTTTG 603
```

seq_name: gb_sts:G30204

seq_documentation_block:
```
LOCUS       G30204           485 bp    DNA        linear   STS 05-OCT-1996
DEFINITION  human STS SHGC-36487, sequence tagged site.
ACCESSION   G30204
VERSION     G30204.1  GI:1593755
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 485)
AUTHORS     Myers,R.M.
JOURNAL     Unpublished (1996)
COMMENT

            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu

            Primer A: ACTTCTTTTTTTCATTGCACAGG
            Primer B: GCCCTTATTTACCGGAAGCT
            STS size: 77
            PCR Profile:
```

Initial incubation: 94 degrees C for 90 seconds

```
Denaturation:     94 degrees C for 15 seconds
Annealing:        62 degrees C for 23 seconds
Polymerization:   72 degrees C for 30 seconds
PCR Cycles:       30
Thermal Cycler:   Perkin Elmer 9600
```

Protocol:
```
  Template:       25 ng
  Primer:         each 1 uM
  dNTPs:          each 200 uM
  Taq Polymerase: 0.05 units/ul
  Total Vol:      10 ul
```

Buffer:
```
  MgCl2:    2.5 mM
  KCl:      50 mM
  Tris-HCl: 20 mM
  pH:       8.3
```

Prepared with primer pairs provided by Sandoz, derived from T90569
-- Washington University/Merck EST sequence.

```
FEATURES             Location/Qualifiers
     source          1..485
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="16"
     STS             69..145
     primer_bind     69..91
     primer_bind     complement(126..145)
BASE COUNT  146 a  99 c  99 g  131 t  10 others
ORIGIN
```

alignment_block:
US-09-801-115-2 x G30204/rev

Align seg 1/1 to reverse of: G30204 from: 1 to: 485

alignment_scores:
```
  Quality:        366.00        Length:   74
  Ratio:          4.946         Gaps:      0
  Percent Similarity: 100.000   Percent Identity: 98.649
```

```
 26 LeuAspIleIleAsnSerLeuValThrThrValPheMetLeuIleValSe 42
    ||||||||||||||||||||||||||||||||||||||||||||||||::
286 CTTGATATTATCAACTCACTGGTAACAACAGTAATTCATCGTCATCGTANC 237

 42 rValLeuAlaLeuIleProGluThrThrLeuThrHisValGlyGlyGlyV 59
    ::||||||||||||||||||||||||||||||||||||||||||||||||
236 TGTGTTGGCACTGATACCAGAAAACCACAACATTGACAGTTGGTGGAGGGG 187

 59 alPheAlaLeuValThrAlaValCysCysLeuAlaAspGlyAlaLeuIle 75
186 TGTTTGCACTTGTGACAGCAGTAGTANCGTCGTTCGCGACGGGCGCCTTATT 137

 76 TyrArgLysLeuLeuPheAsnProSerGlyProTyrGlnLysLysProVa 92
    |||||||||||||||||||||||||||||||||||||||||||||||||
136 TACCGGAAGCTTCTGTTCAATCCCAGGCGGTCCTTACCAGAAAAAGCCTGT 87

 92 lHisGluLysLysGluValLeu 99
    |||||||||||||||||||||
 86 GCATGAAAAAAAAGAAGTTTTG 65
```

seq_name: gb_pr:AF145216

seq_documentation_block:
```
LOCUS       AF145216          593 bp    mRNA       linear   PRI 07-SEP-2000
DEFINITION  Homo sapiens chemokine-like factor 4 (CKLF4) mRNA, complete cds,
            alternatively spliced.
ACCESSION   AF145216
VERSION     AF145216.2  GI:9989694
KEYWORDS    .
```

SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 593)
  AUTHORS   Han,W.L., Gu,W.F., Li,Y., Zhang,Y.M., Di,C.H., Song,Q.S. and
            Ma,D.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-APR-1999) Immunology, Beijing Medical University, Xue
            Yuan Road, Beijing 100083, China
REFERENCE   2 (bases 1 to 593)
  AUTHORS   Han,W.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-SEP-2000) Immunology, Peking University Health
            Science Center, Xueyuan Road 38, Beijing 100083, China
  REMARK   Sequence update by submitter
COMMENT   On Sep 7, 2000 this sequence version replaced gi:6625671.
FEATURES   Location/Qualifiers
     source   1..593
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="U937"
                     /note="from PHA stimulated cells"
     gene   1..593
                     /gene="CKLF4"
                     /note="UCK"
     CDS   148..510
                     /gene="CKLF4"
                     /note="UCK-4; alternatively spliced"
                     /codon_start=1
                     /product="chemokine-like factor 4"
                     /protein_id="AAF19350.1"
                     /db_xref="GI:6625672"
                     /translation="MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEP
                     YIVITGFEVTVILFFILLYVLRLDRLMKWLFWPLLVFALVTAVCCLADGALLYRKLLF
                     NPSGPYQKKPVHEKKEVL"
ORIGIN
BASE COUNT   166 a   120 c   140 g   167 t

alignment_block:
US-09-801-115-2 x AF145216

Align seg 1/1 to: AF145216 from: 1 to: 593

alignment_scores:
  Quality: 339.50   Length: 126
  Ratio: 4.297   Gaps: 2
  Percent Similarity: 62.698   Percent Identity: 57.937

```
 1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
   ||||||||||||||||||||||||||||||||||||||||||||||||
 1 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGT 197

17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
   ||||||||||||||||||||||||||||||||                :::
198 GAAAGGCCACGTGAAGAGTGCTGCGGCTGGCACTA.               231

34 hrThrValPheMetLeuLeuIleValSerValLeuAlaLeuIleProGluThr 50
   ||||||  :::                                     ::::||
232 ..ACTGTGACACTGTAGACCTTTTTATCATCGCACAAGCCCCTGAACCA  279

51 ThrThrLeuThrValGlyGlyGly.......................... 58
   :: :|||
280 TATATTGTTTATCACTGGATTGAAGTCACCGTTATCTTATTTTTTCATACT 329

58 ................................................. 58
330 TTTATATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTT 379

59 ......ValPheAlaLeuValThrAlaValCysCysLeuAlaAspGlyAla 73
      ||||||||||||||||||||||||||||||||||||
380 TGCTTGTGTTTGTTTGCACTTGTGACAACAGTATGCTGTCGTTGCCGACGGGGCC 429
```

LOCUS   AF135381   434 bp   mRNA   linear   PRI 07-SEP-2000
DEFINITION Homo sapiens chemokine-like factor 3 (CKLF3) mRNA, complete cds,
            alternatively spliced.
ACCESSION   AF135381
VERSION   AF135381.2 GI:9989693
KEYWORDS   .
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 434)
  AUTHORS   Han,W.L., Gu,W., Li,Y., Zhang,Y., Song,Q., Di,C. and Ma,D.
  TITLE   Direct Submission
  JOURNAL   Submitted (17-MAR-1999) Immunology, Beijing Medical University, Xue
            Yuan Road, Beijing 100083, China
REFERENCE   2 (bases 1 to 434)
  AUTHORS   Han,W.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-SEP-2000) Immunology, Peking University Health
            Science Center, Xueyuan Road 38, Beijing 100083, China
  REMARK   Sequence update by submitter
COMMENT   On Sep 7, 2000 this sequence version replaced gi:6630855.
FEATURES   Location/Qualifiers
     source   1..434
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="U937"
                     /note="from PHA stimulated cells"
     gene   1..434
                     /gene="CKLF3"
                     /note="UCK"
     CDS   148..351
                     /gene="CKLF3"
                     /note="UCK-3; alternatively spliced"
                     /codon_start=1
                     /product="chemokine-like factor 3"
                     /protein_id="AAF19600.1"
                     /db_xref="GI:6630856"
                     /translation="MDNVQPKIKHRPFCFSVKGHVKMLRLVFALVTAVCCLADGALIY
                     RKLLRNPSGPYQKKPVHEKKEVL"

```
74 LeuIleTyrArgLysLeuLeuPheAsnProSerGlyProTyrGlnLysGly 90
   ||||||||||||||||||||||||||||||||||||||||||||||||||
430 CTTAATTACCGGAAGCTTCTGTTCAATCCAGCGGGTCCTTACCAGAAAA 479

90 sProValHisGluLysLysGluValLeu 99
   |||||||||||||||||||||||||||
480 GCCTGTGCATGAAAAAAAAAGAAGTTTG 507
```

seq_name: gb_pr:AF135381

seq_documentation_block:

ORIGIN
BASE COUNT   127 a   88 c   116 g   103 t

alignment_block:
US-09-801-115-2 x AF135381

Align seg 1/1 to: AF135381 from: 1 to: 434

alignment_scores:
  Quality: 331.00   Length: 99
  Ratio: 4.940   Gaps: 1
  Percent Similarity: 67.677   Percent Identity: 67.677

```
 1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
   |||||||||||||||||||||||||||||||||||||||||||||||
148 ATGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGT 197

17 lLysGlyHisValLysMetLeuArgLeuAspIleIleAsnSerLeuValT 34
   |||||||||||||||||||||||||||
198 GAAAGGCCACGTGAAGAGTGCTGCGGCTG...................... 225
```

Left column

```
 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr
    |||||||||||||||||||||||||||||||||||||||||||||||||
225
```

```
225 ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy 67
    |||||||||||||||||||||||||||||||||
 51         GTGTTTGCACTTGTGACAGCAGTATG 67
```

```
226 ..................................................
                                                   251
 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnProS 84
    |||||||||||||||||||||||||||||||||||||||||
252 CTGTCTTGCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAACAAT 301
```

```
 84 erGlyProTyrGlnLysLysProValHisGluLysGluVaLLeu 99
    |||||||||||||||||||||||||||||||||||
302 GCGGTCCTTACCAGAAAAAGCCTGTGCATGAAAAAAAAAGAAGTTTTG 348
```

seq_name: gb_pat:AX079435

LOCUS   AX079435   495 bp   DNA   linear   PAT 22-FEB-2001
DEFINITION Sequence 179 from Patent WO0107611.
ACCESSION   AX079435
VERSION   AX079435.1   GI:13158995
KEYWORDS   .
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 495)
  AUTHORS   Baker,K.P., Goddard,A. and Wood,W.I.
  TITLE   Human polypeptides and methods for the use thereof
  JOURNAL   Patent: WO 0107611-A 179 01-FEB-2001;
            Genentech, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..495
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT   119 a   103 c   127 g   142 t   4 others
ORIGIN

alignment_scores:
    Quality: 318.50   Length: 124
    Ratio: 4.550   Gaps: 1
Percent Similarity: 56.452   Percent Identity: 56.452

alignment_block:
US-09-801-115-2 x AX079435

Align seg 1/1   to: AX079435   from: 1   to: 495

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
    |||||||||||||||||||||||||||||||||||||||||||
  1 ATGGATAACGTGCAGCCCAAAAATAAAACATCGCCCTTCTGCTTCAGTGT 173
```

```
 17 LLysGlyHisValLLysMetLeuArg.................. 25
    |||||||||||||||||||||
174 GAAAGGCCACGTGAAGATGCTGCGGCGTGGCACTAACTGNGACATCTATGA 223
```

```
 25 .................................................. 25
224 CCTTTTTTATNATCGCACAAGCCCCTGAACCANTATNATTGTTATCACNTGGA 273
```

```
 25 .................................................. 25
274 TTTGAAGTCACCGTTATCTTTATTTTCATACTTTTATTATGTACTCAGACT 323
```

```
 26 ............................LeuAspIleIleAsnS 31
                                  |||||||||||||||
324 TGATCGATTAATGAAGTGGTTATTTTTGGCCTTTGCTTTGATATTATCAACT 373
```

```
 31 erLeuValThrThrValPheMetLeuIleValSerValLeuAlaLeuIle 47
```

Right column

```
374 CACTGGTAACAACACAGTATTCATTCATCGTATCGTGTTGGCACTGATA 423
    ||||||||||||||||||||||||||||||||||||||||||||||||
```

```
 48 ProGluThrThrThrLeuThrValGlyGlyValPheAlaLeuValTh 64
    |||||||||||||||||||||||||||||||||||||
424 CCAGAAACCACCACACATTGACACAGTTGGTGGAGGGGTGTTTGCACTTGTGAC 473
```

```
 64 rAlaValCysCysLeuAlaAsp 71
    ||||||||||||
474 AGCAGTATGCCTGTNTTGCCGAC 495
```

seq_name: gb_ro:AF253064

LOCUS   AF253064   523 bp   mRNA   linear   ROD 27-NOV-2001
DEFINITION Rattus norvegicus chemokine-like factor 1 (CKLF1) mRNA, complete
            cds.
ACCESSION   AF253064
VERSION   AF253064.2   GI:17105378
KEYWORDS   .
SOURCE   Norway rat.
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 523)
  AUTHORS   Lou,Y., Li,Y., Han,W., Song,Q., Zhang,Y., Di,C. and Ma,D.
  TITLE   Direct Submission
  JOURNAL   Submitted (06-APR-2000) Immunology, Beijing Medical University,
            Xueyuan Road, Beijing 100083, China
REFERENCE   2 (bases 1 to 523)
  AUTHORS   Lou,Y., Li,Y., Han,W., Song,Q., Zhang,Y., Di,C. and Ma,D.
  TITLE   Direct Submission
  JOURNAL   Submitted (27-NOV-2001) Immunology, Beijing Medical University,
            Xueyuan Road, Beijing 100083, China
REMARK   Sequence update by submitter
COMMENT   On Nov 27, 2001 this sequence version replaced gi:7769680.
FEATURES   Location/Qualifiers
     source   1..523
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /tissue_type="liver"
     gene   1..523
                     /gene="CKLF1"
     CDS   167..463
                     /gene="CKLF1"
                     /note="cytokine; similar to Homo sapiens CKLF1"
                     /codon_start=1
                     /product="chemokine-like factor 1"
                     /protein_id="AAF69502.1"
                     /db_xref="GI:7769681"
                     /translation="MDSPQKVVDHQPFCLSLKCFVXTLRIDVINSVVTTLFMLIVSVS
                     ALIPETSTMIMVGVFGFLFVICITVADCALMCQKILRFRHGPYQNRSATDVDDS"
BASE COUNT   134 a   107 c   133 g   149 t
ORIGIN

alignment_scores:
    Quality: 275.00   Length: 90
    Ratio: 3.667   Gaps: 0
Percent Similarity: 83.333   Percent Identity: 58.889

alignment_block:
US-09-801-115-2 x AF253064

Align seg 1/1   to: AF253064   from: 1   to: 523

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
    |||||:||||||||||||||||||||:|||:|::||||||||||||||
167 ATGGACTCTCCACAGAAGGTCGTGAGACCATCAGCCCTTCTGCCTCAGTCT 216
```

```
 17 LLysGlyHisValLLysMetLeuArgLeuAspIleIleAsnSerLeuValTr 34
```

seq_name: gb_pat:AX330610

```
217 GAAATGCTTGTGAAGACGCTCGCGGTTGGAGATTGGAATTATCAACTCGGTTGGTAA 266
         :::||||||||||||||||||||||||||||||||||||:::||||||||:||||
 34 hrThrValPheMetLeuIleValSerValLeuAlaLeuIleProGluThr 50
267 CAACACTATTCATCCTTATTGTGTCTGTCGTCGTCCTCTAATACCAGAAACC 316
       ||||||||||||||||||||||||||||||||||||||||||||:::||||
 51 ThrThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCy 67
317 TCAACAATGATAATATGGTTGGAGGGGTTGTTTCTTGACAGTAAATATG 366
       ::|||||||||||:|:::||||||||||||||||||||:||||||||
 67 sCysLeuAlaAspGlyAlaLeuIleTyrArgLysLeuPheAsnPros 84
367 TACTGTTGCTGATTGTGCCTTATGTGCCAGAAAACTCCGGTTCGTCCAC 416
       ||||||||||||||||||||||||||||||||||||||||||||||
 84 erGlyProTyrGlnLysLys 90
417 ATGGACCTTATCAGAACAGG 436
       |||||||||||::::::
```

seq_documentation_block:
LOCUS   AX330610   207 bp   DNA   linear   PAT 09-JAN-2002
DEFINITION Sequence 1119 from Patent WO0194629.
ACCESSION   AX330610
VERSION   AX330610.1 GI:18103588
KEYWORDS   .
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE   Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 1119 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
     source   1..207
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT   70 a   43 c   39 g   54 t   1 others
ORIGIN

alignment_block:
US-09-801-115-2 x AX330610/rev   ..

alignment_scores:
    Quality: 245.00   Length: 48
    Ratio: 5.213   Gaps: 0
    Percent Similarity: 97.917 Percent Identity: 97.917

Align seg 1/1 to reverse of: AX330610 from: 1 to: 207

```
 52 ThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCysCy 68
205 ACAATTGACACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTG 156
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 68 sLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnProSerG 85
155 TCTTGCCGACAGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCG 106
    |||||||||||||||||||||||||||||||||||||||||||||||||
 85 lyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
105 GTCCTTACCAGNAAAAGCCTGTGCATGAAAAAAAGAAGTTTTG 62
    |||||||||||||||||||||||||||||||||||||||||
```

seq_documentation_block:
LOCUS   AX330787
DEFINITION Sequence 1296 from Patent WO0194629.

ACCESSION   AX330787
VERSION   AX330787.1 GI:18103765
SOURCE   .
KEYWORDS   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE   Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 1296 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
     source   1..207
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT   70 a   43 c   39 g   54 t   1 others
ORIGIN

seq_name: gb_pat:AX331008

alignment_scores:
    Quality: 245.00   Length: 48
    Ratio: 5.213
    Percent Similarity: 97.917 Percent Identity: 97.917

Align seg 1/1 to reverse of: AX330787 from: 1 to: 207

```
 68 sLeuAlaAspGlyAlaLeuIleTyrArgLysLeuLeuPheAsnProSerG 85
205 ACAATTGACACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTG 156
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 52 ThrLeuThrValGlyGlyValPheAlaLeuValThrAlaValCysCy 68
155 TCTTGCCGACAGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCG 106
    |||||||||||||||||||||||||||||||||||||||||||||||||
 85 lyProTyrGlnLysLysProValHisGluLysLysGluValLeu 99
105 GTCCTTACCAGNAAAAGCCTGTGCATGAAAAAAAGAAGTTTTG 62
    |||||||||||||||||||||||||||||||||||||||||
```

alignment_block:
US-09-801-115-2 x AX331008/rev   ..

alignment_scores:
    Quality: 245.00   Length: 48

seq_documentation_block:
LOCUS   AX331008   207 bp   DNA   linear   PAT 09-JAN-2002
DEFINITION Sequence 1517 from Patent WO0194629.
ACCESSION   AX331008
VERSION   AX331008.1 GI:18121642

Percent Similarity: 97.917

alignment_block:
US-09-801-115-2 x AX331008/rev

Ratio: 5.213   Gaps: 0
Percent Similarity: 97.917   Percent Identity: 97.917

Align seq 1/1 to reverse of: AX331008 from: 1 to: 207 ..

```
 52 ThrLeuThrValGlyGlyValGlyValPheAlaLeuValThrAlaValCysCy  68
    |||||||||||||||||||||||||||||||||||||||||||||||||||
205 ACATTGACAGTTGGTGGGGTGGGGGTGTTTGCACTTGTGACAGCAGTATGCTG 156

 68 sLeuAlaAspGlyAlaLeuIleThrTyrArgLysLeuLeuPheAsnProSerG  85
    |||||||||||||||||||||||||||||||||||||||||||||||||||
155 TCTTGCCGACGGGGCCCTTATTTACCCGGAAGCTTCTGTTCAATCCCAGCG 106

 85 lyProTyrGlnLysLysProValHisGluLysGlyGluValLeu  99
    |||||||||||||||||||||||||||||||||
105 GTCCTTACCAGNAAAAGCCTGTGCATGAAAAAAAAGAAGTTTG 62
```

seq_name: gb_ro:AF253065

alignment_scores:
  Quality: 238.50   Length: 143
  Ratio: 3.180   Gaps: 1
  Percent Similarity: 52.448   Percent Identity: 37.063

---

alignment_block:
US-09-801-115-2 x AF253065

Align seq 1/1 to: AF253065 from: 1 to: 682 ..

```
  1 MetAspAsnValGlnProLysIleLysHisArgProPheCysPheSerVa 17
    ||||||||:::   |||   :::||||||||||||||:::||||||:::::
167 ATGGACTCTCCACAGAAGGTCGTAGAACCATCAGCCCTTCTGCCTCAGTCT 216

 17 lLysGlyHisValLysMetLeuArg.................... 25
    :||   ||||||||||:|:||||||
217 GAAAATGCTTTGTTGAAGACGCGCTGCGGGTTGGTGGTAACTGTGGCCTCTATGA 266

 25 ..................... 25
                           
267 TCTTTTCATTGTGGCACAAGCCCCTGAACCGTATATCGTTATCACTGGG 316

 25 ..................... 25
                           
317 TTTGAAGTGCACCATCATTCTCTTTCTCATAGCCTTATATATGTGCAGTCT 366

 26 ..................... LeuAspIleIleAsnS 31
                            ||:||||||||||||
367 TGACAGACAATGAGATCTTTCTTTTGGCCTTTGCTTGAGTTATCAACT 416

 31 erLeuValThrHrvalPheMetLeuIleValSerValLeuAlaLeuIle 47
    |||:::|||||||:|||||||||||||||||:|||||||||||||||||
417 CGGTGGTGAACAACACTAATCAAGCTTATTGTGTCTGTCGGCTCTAATA 466

 48 ProGluThrThrThrLeuThrValGlyGlyValGlyValPheAlaLeuValTh 64
    |||||||||||||||||||||||||||||||||::|||||||||||||||||
467 CCAGAAAACCTCAACAATGATAATGGTTGGAGGGGTGTTTGGTTTCCTGAC 516

 64 rAlaValCysCysLeuAlaAspGlyAlaLeuIleThrTyrArgLysLeuLeuP 81
    |:::||||||||:|||||||:::||||||:::||||||||||||:||||||
517 AGTAATAATGTACTTGTTGCCTGAGTTGTGCCCTTATGTGCCAGAAAACTCCGGT 566

 81 heAsnProSerGlyProTyrGlnLysLys 90
    ||:::|||   |||||||||||||:::::
567 TTCGTCCACATGGACCTTATCAGAAACAGG 595
```